

Generatywna Sztuczna Inteligencja

Kreatywność

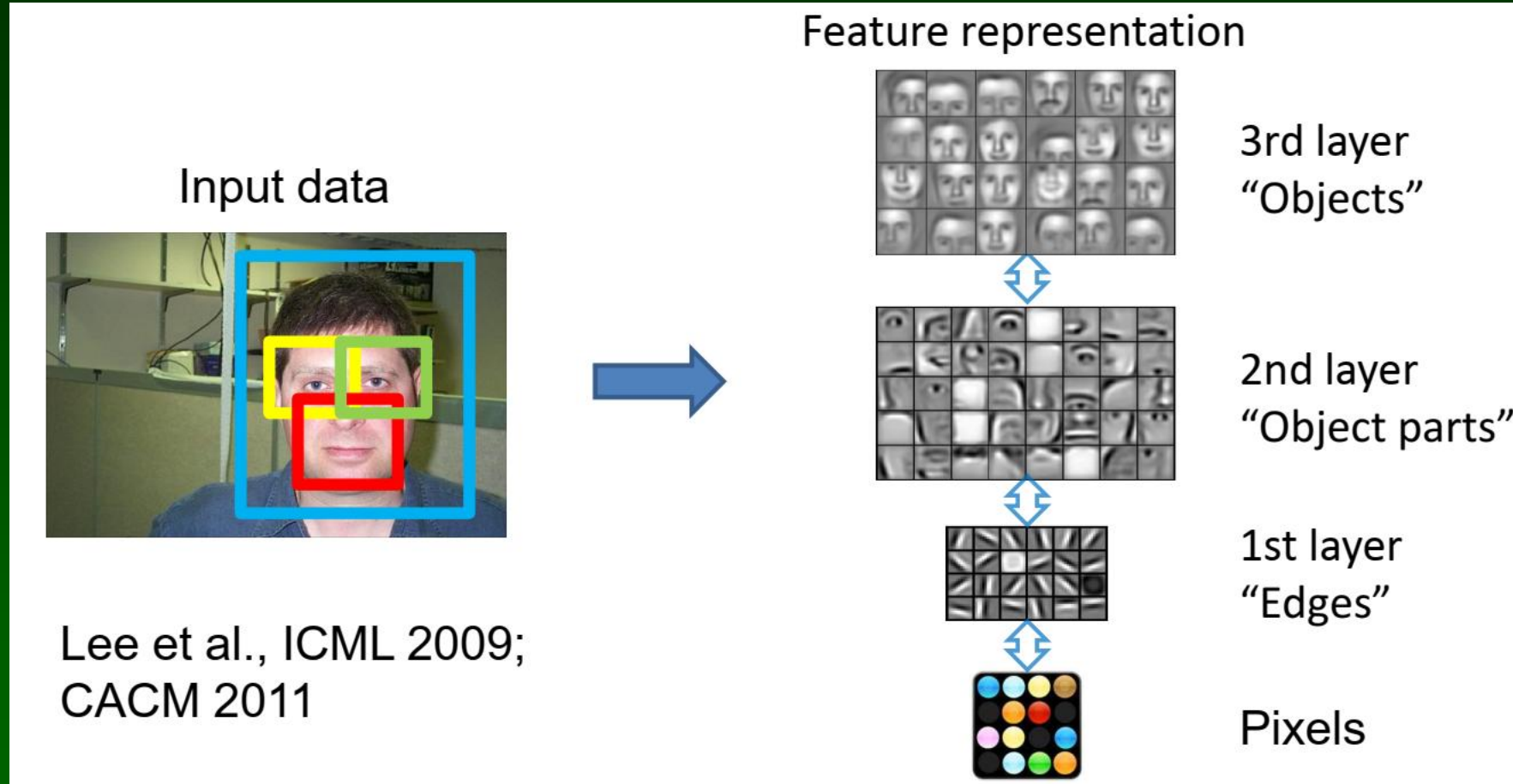


Włodzisław Duch

Katedra Informatyki Stosowanej UMK

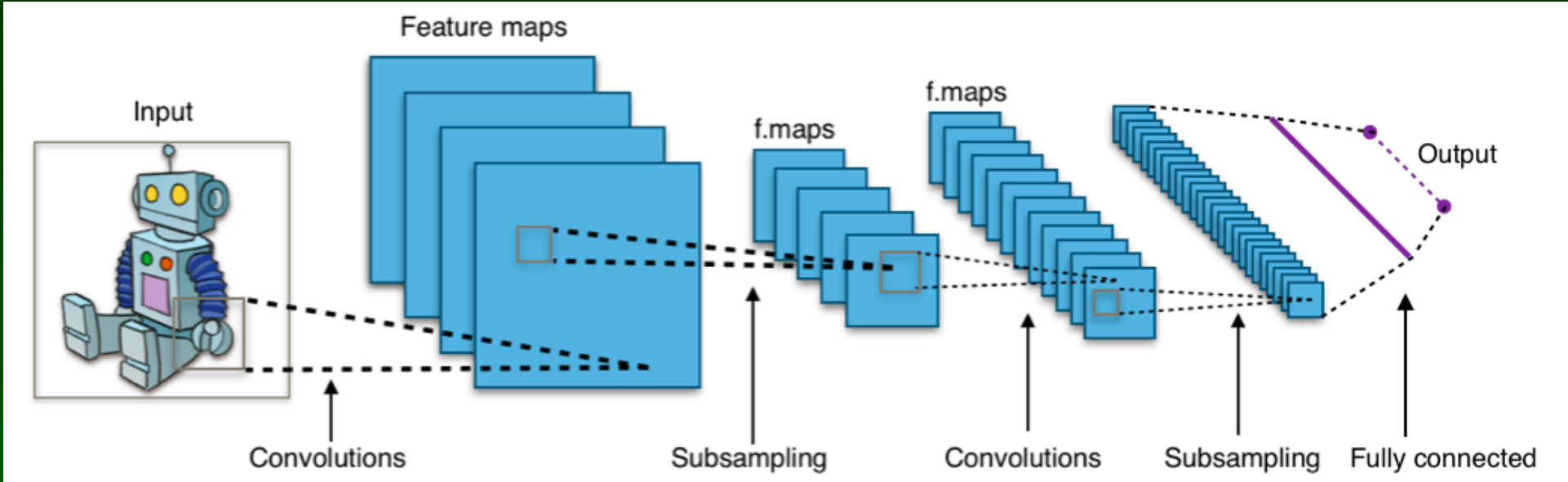
Google: Włodzisław Duch

Methods: NN for images



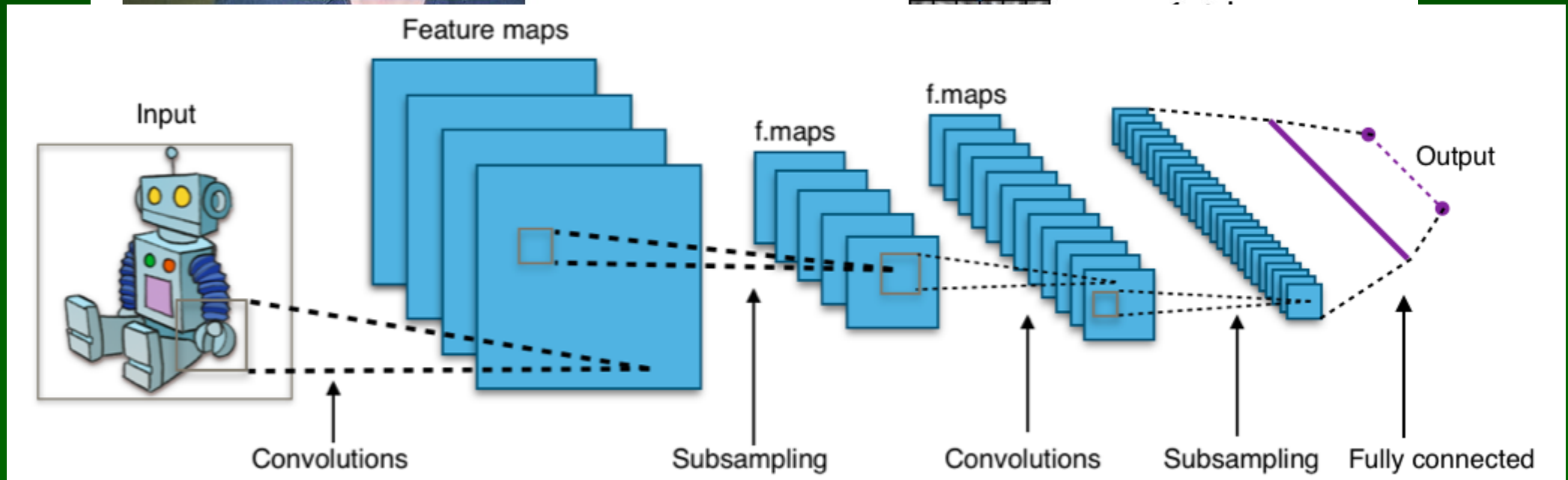
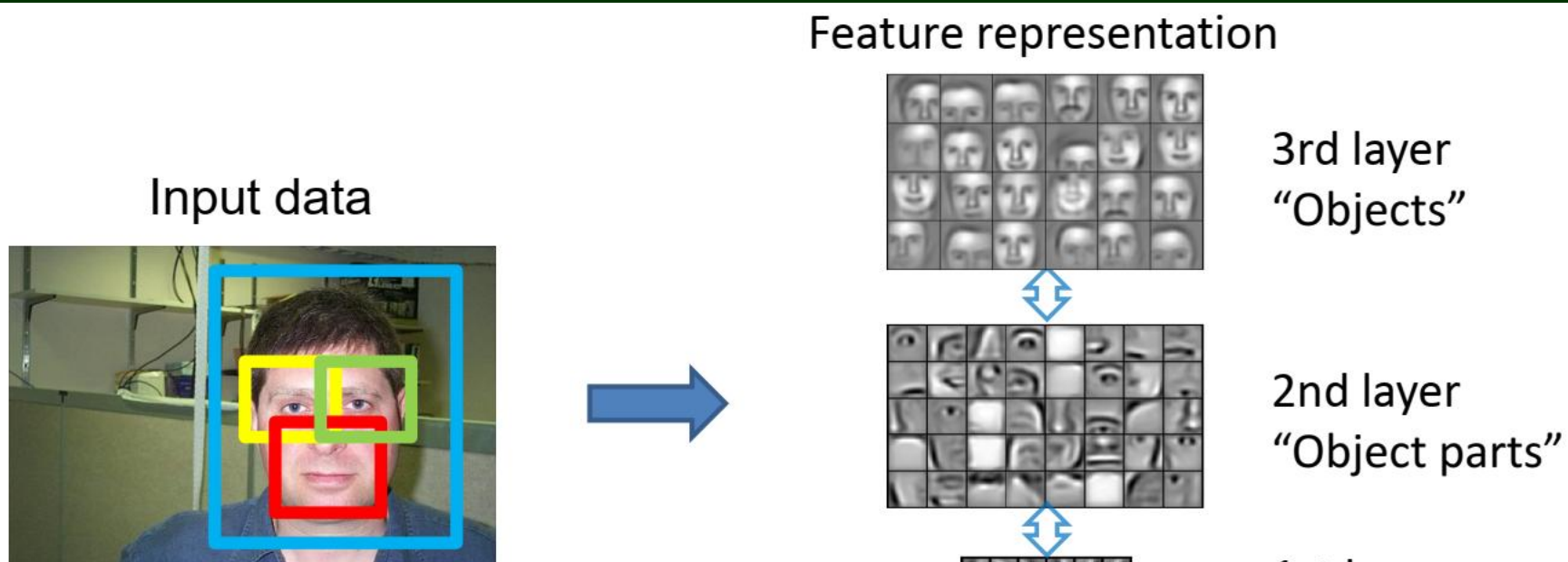
IBM an [introduction to deep learning](#)

Methods: Deep NN for images



Small patches of image (ex. 5x5 pixels) are filtered through receptive fields in convolution operation, creating elementary features, combined together using sampling, and again filtered to generate more complex features. Finally generated features are analyzed to determine image category.

Methods: NN for images



Superresolution

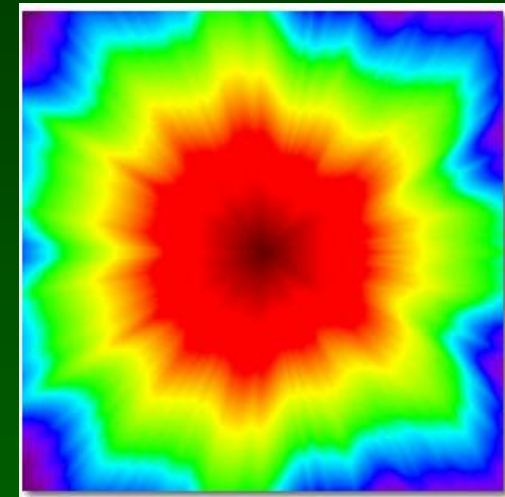
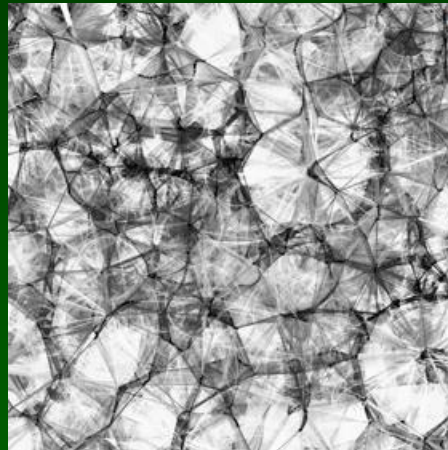


Generative art

„Generative art” w Wikipedii, początki AI w sztuce

http://en.wikipedia.org/wiki/Generative_art

<http://www.random-art.org>



<http://www.solaas.com.ar/dreamlines/p5/>

<http://www.jhllabs.com/java/art.html>

Yuki Terai singing animated avatar, first digital idol from Japan!



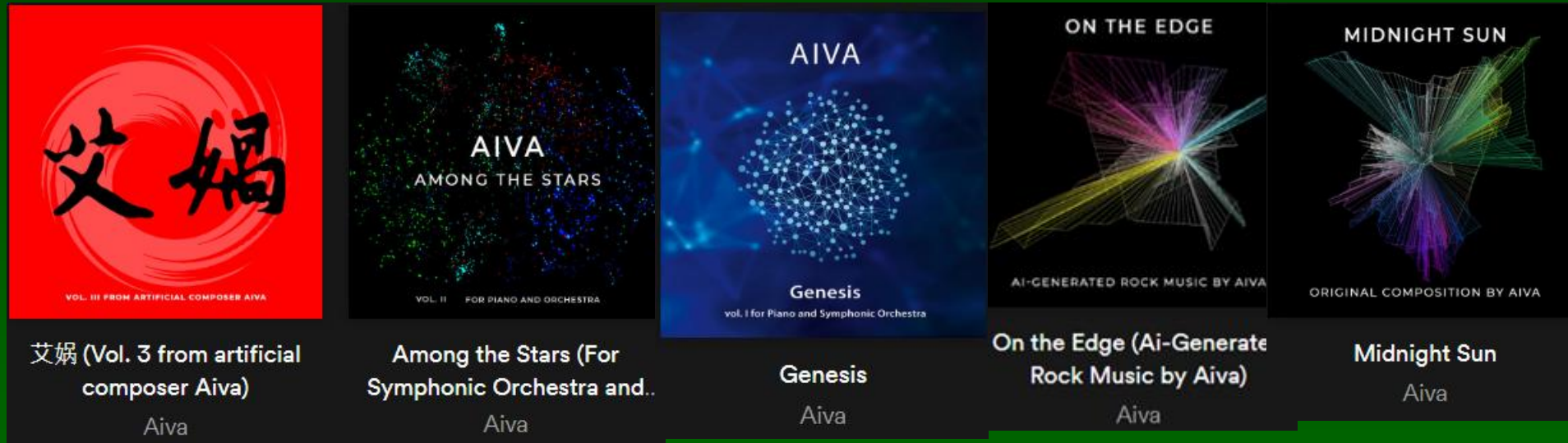
AI Virtual Artist

[AIVA](#) – AI Virtual Artist, member of author's rights society ([SACEM](#)), 206 works.

[AIVA YouTube](#) channel, Youtube „[Letz make it happen](#)“, Op. 23

[SoundCloud channel](#)

[Spotify](#) and [Apple](#) channel



Aaron Painter

Harold Cohen has asked about minimal requirements that should be fulfilled to call something art painting, and created in 1973 [Aaron](#) robotic artist.



Aaron examples



Meeting On Gauguin's Beach,



Aaron, with Decorative Panel, 1988
Oil 1992

Basic principles of composition + randomization. [See more here.](#)

Aaron in museums



Liberty & Friends, 1985
Ink



Mural, Capitol
Children Museum 1980

Nowy Rembrandt

Next Rembrandt Project.

- All known 346 Rembrandt van Rijn portraits were scanned in 3D, about ~150 GB.
- Learning to paint in Rembrandt style, detailed movement of brush, using deep neural networks.
- Creation of new portrait in Rembrandt style— 500 hours of image rendering, 148 mln pixels in 3D.
- Printing this with 13 layers of paint.

i



Deep Art



<https://deepart.io/latest/> Deep Neural Networks, Deep Dream,
LA Gatys, AS Ecker, M Bethge, A Neural Algorithm of Artistic Style (2015)

Examples of Deep Art



<https://deepart.io/latest/> Deep Art,
LA Gatys, AS Ecker, M Bethge, A Neural Algorithm of Artistic Style (2015)



[Google Deep Dream/Deep Style i Generator](#),
LA Gatys, AS Ecker, M Bethge, A Neural Algorithm of Artistic Style (2015)

Generative music

„Popcorn”, Hot Butter przebój z 1969,
Stworzony dzięki analizie statystycznej różnych melodii



Muzycy uczą się słuchając i analizując różne utwory.
Czemu komputery nie mogłyby też się tak uczyć? Mając dostęp do całej muzyki na świecie?

Muzyka algorytmiczna: John Cage, Iannis Xenakis ...

http://en.wikipedia.org/wiki/Algorithmic_music

Lucasfilm Games' gra Ballblazer (1982) używała komputerowej improwizacji jazzowej, Xbox improwizuje na zadane tematy.

Muzyka generatywna: Brian Eno już w 1975 robił takie eksperymenty!

http://en.wikipedia.org/wiki/Generative_music

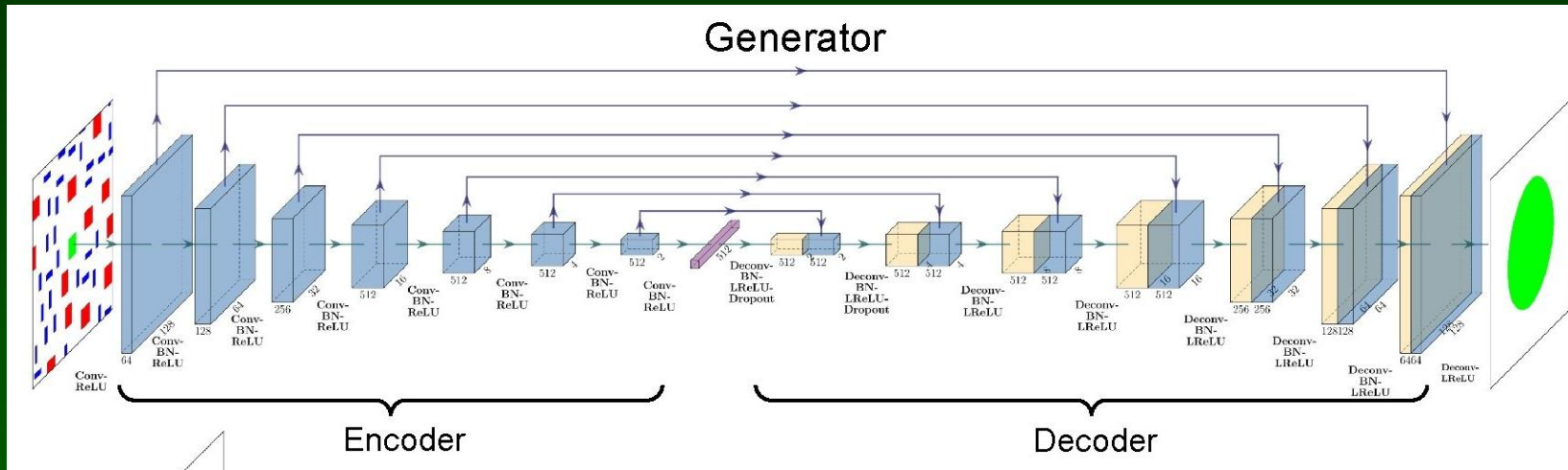
Przykłady wczesnej muzyki generatywnej:

<http://jmusic.ci.qut.edu.au/jmMusic.html>



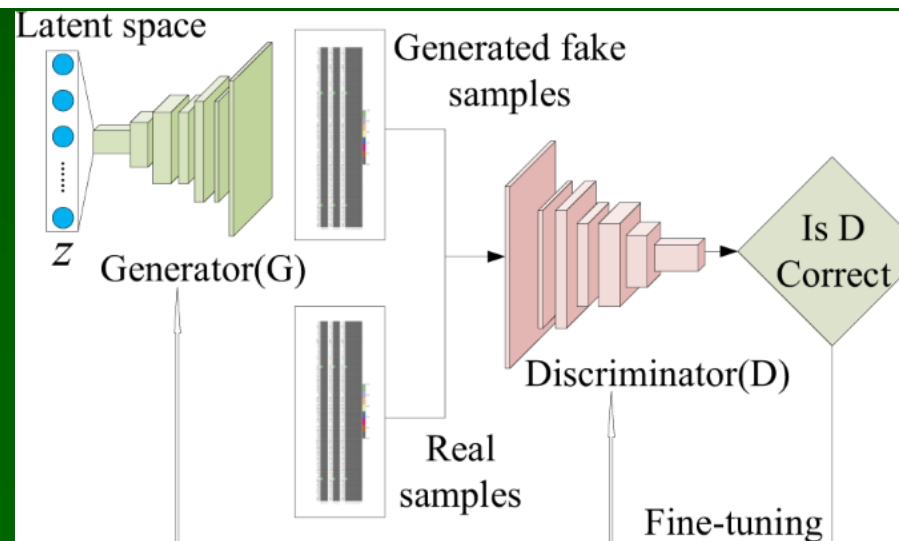
Obrazy i kreatywność

Autoenkodery – kompresja i dekompresja informacji, tworzy przestrzeń ukrytych parametrów.



Generative Adversarial Networks:
Generuj obrazy (G), porównuj z rzeczywistymi (D), zmieniaj parametry by stawały się coraz bardziej realistyczne.

[Przykłady na Youtube](#), [Dall-e2](#)



From images to music

Can music be composed converting images?

Example: finish “metacomposer” Lauri Gröhn:

Synesthesia Software Music

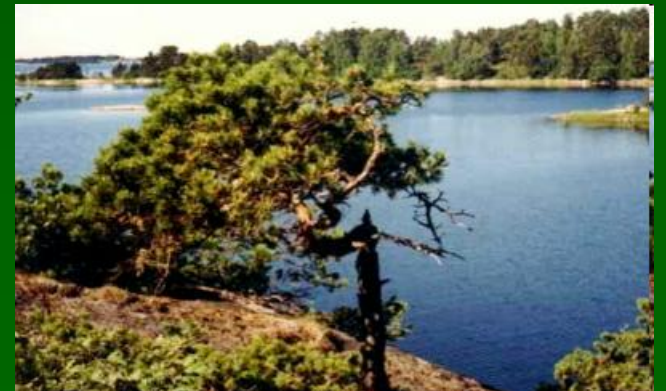
„I want to show that in future computers may compose as well as people”.



Turku Archipelago, Fog, Sakura – interesting compositions!

Can computer really „feel” music?

Greta ist trying...



Flow Machines

by Sony CSL

<http://www.flow-machines.com>

Software się uczy, to co zaprogramowane
to tylko motywacja by się uczyć.

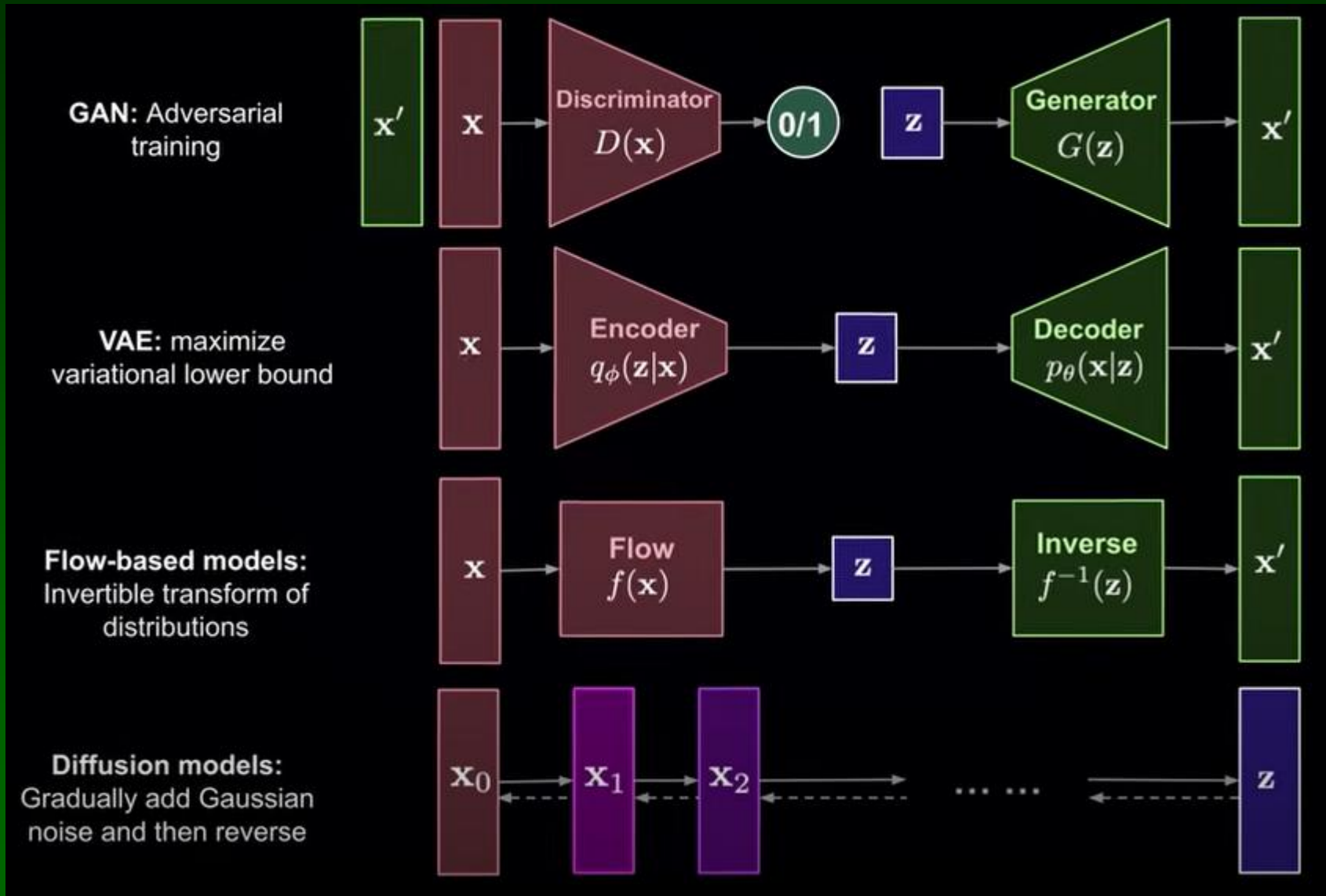
[Learning to Create](#) (Lrn2Cre8)

Daddy's Car: a song composed by Artificial Intelligence - in the style of the Beatles



flowmachines
flowmachines

Kompresja i generatywna AI



Wiele metod kompresji.

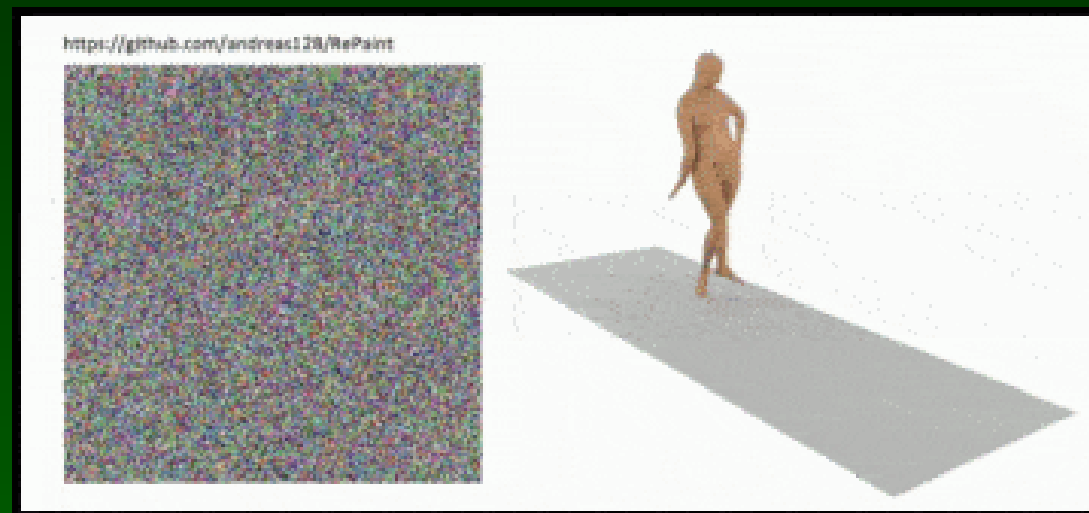
Po kompresji informacja jest znacznie prostsza, mniej dokładna, mamy fragmenty układanki.

Prawdopodobieństwa, a nie całe obrazy czy zapamiętane teksty.

Polecenia (prompty) pozwalają na ułożenie skojarzonych z nimi fragmentów w nową sensowną całość.

Przyszło mi do głowy ...

Wiele modeli tekst-obraz, tekst-wideo, używa ukrytych modeli dyfuzyjnych:
Stability AI Stable Diffusion, OpenAI DALL·E 2, Google Imagen, GLIDE, Midjourney.
RectifiedHR (3/2025) tworzy obrazy w 1/10 sekundy!



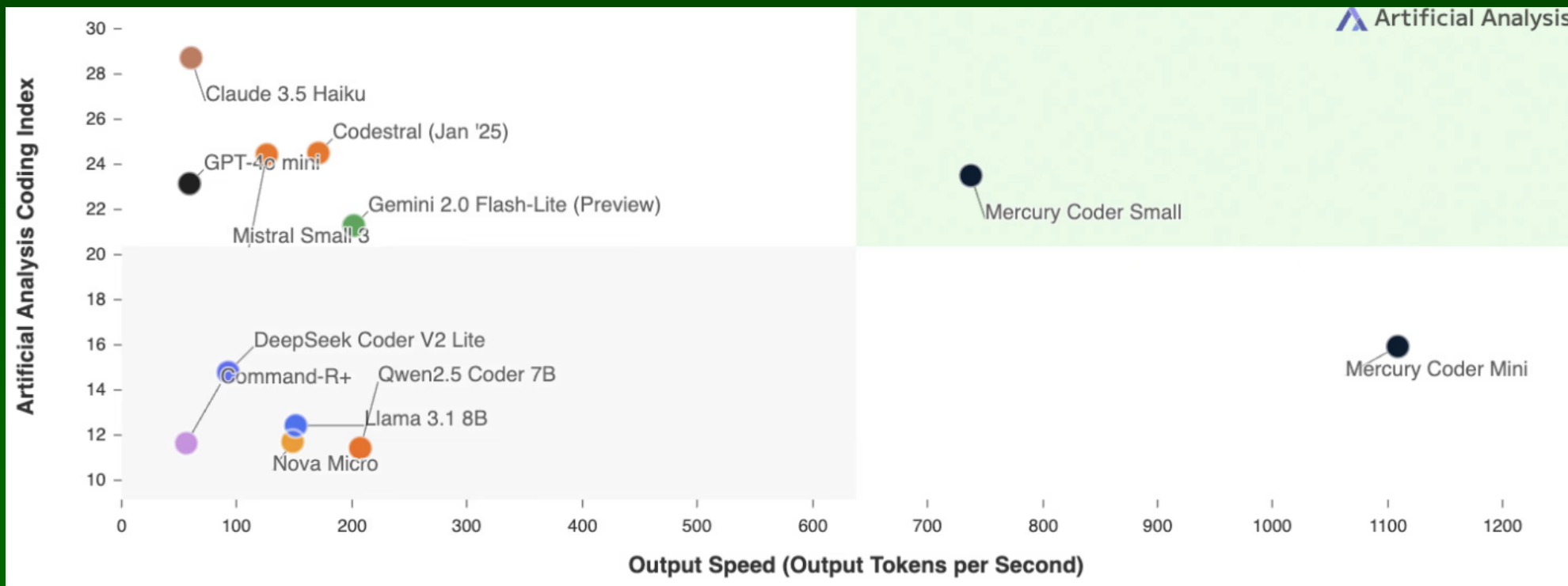
Modele dyfuzyjne mogą tworzyć całe zdania jednocześnie, zamiast tworzyć kolejne słowa.

Nie, S. et al. (2025). Large Language Diffusion Models. [arXiv:2502.09992](https://arxiv.org/abs/2502.09992)

LLaDA (Large Language Diffusion with mAsking)

Dyfuzyjne LLMs

- [Inception Labs](#) Mercury, pierwszy dyfuzyjny model językowy na skalę komercyjną.
- Rodzina dyfuzyjnych dużych modeli językowych (dLLM) Mercury, nowa generacja LLM, która przesuwa granicę szybkiego generowania wysokiej jakości tekstu. Generuje ponad 1000 tokenów/s na komputerach NVIDIA H100.
- Model generowania kodu, Mercury Coder, jest dostępny do testowania.
- Takie modele umożliwią głębokie rozumowanie i odpowiedzi bez żadnych opóźnień.

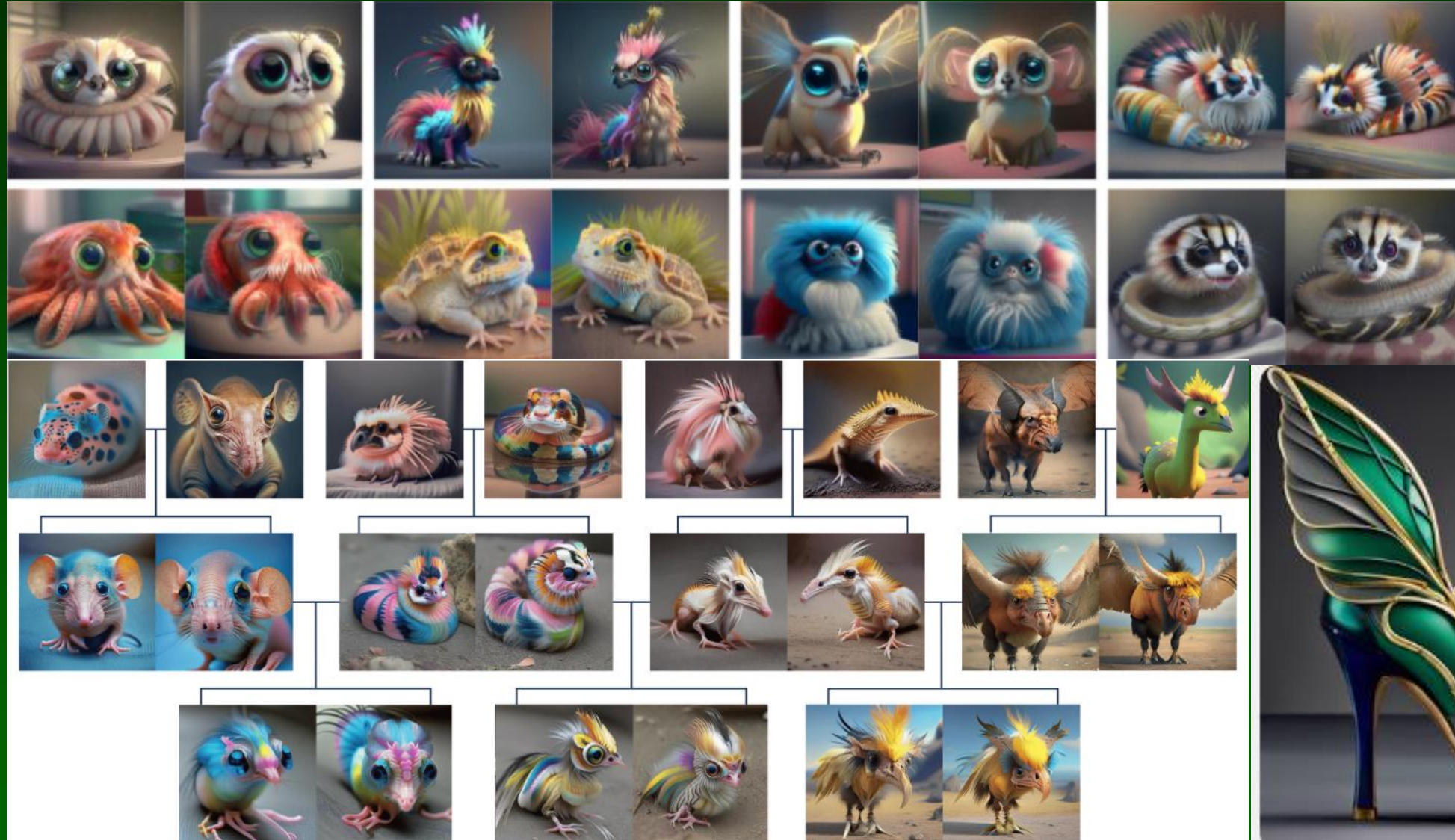


Generatywna AI: Obrazy, wideo, muzyka ...

Luty 2024: SORA, nowy system tekst => realistyczne wideo i obrazy, doskonałe połączenie. Aktorzy i scenarzyści w Hollywood protestują. Powstają tuziny programów GenAI, tworząc tekst, programy, obrazy, grafikę, wideo, kopiując głos ludzki, komponując i wykonując muzykę ...
Zobaczyć znaczy uwierzyć? Już nie. Trzeba wrócić do sprawdzonych źródeł informacji.

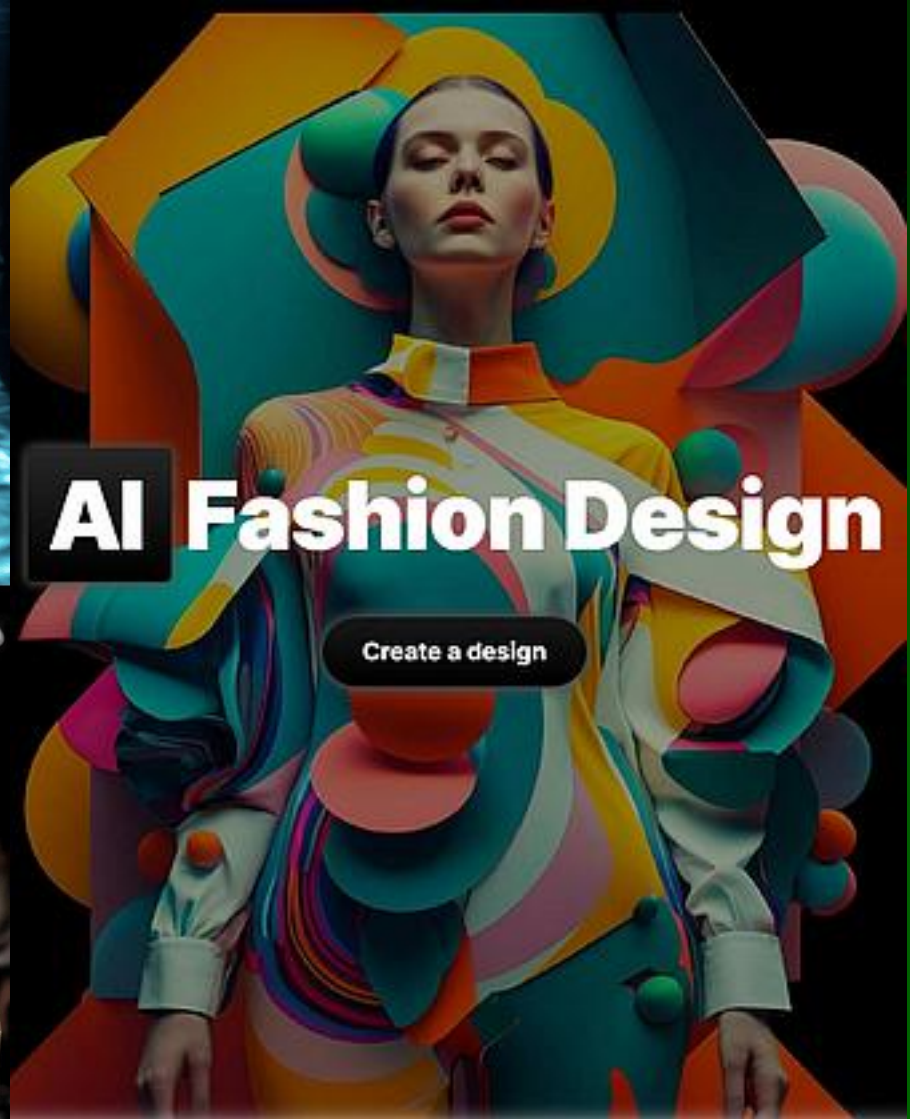


Projektowanie: zabawki i buty



Shoes

Design: tydzień mody AI (NY, 4/2023)



Generative AI for digital art and music

"Edmond de Belamy" – portret namalowany przez AI sprzedany na aukcji Christie w 2018 roku za 432 500\$.



"Théâtre D'opéra Spatial", zwycięzca dorocznego konkursu Colorado State Fair's w 2022 roku.



'A.I. God. Portrait of Alan Turing', namalowany przez robota Ai-Da sprzedano za 1.32-mln \$ na aukcji Sotheby w 2024 r. Część 5-panelli, pokazanych na szczycie United Nations na temat A.I. w Genewie.

AIVA należy do stowarzyszenia kompozytorów SACEM w Francji. Suno, Udio, Soundful, Mubert – tworzą muzykę.



Halucynacje: ludzie i LLMy

Cena kreatywności: kompresja informacji w sieciach neuronowych umożliwia nowe skojarzenia, ale **część będzie zawsze błędna!**

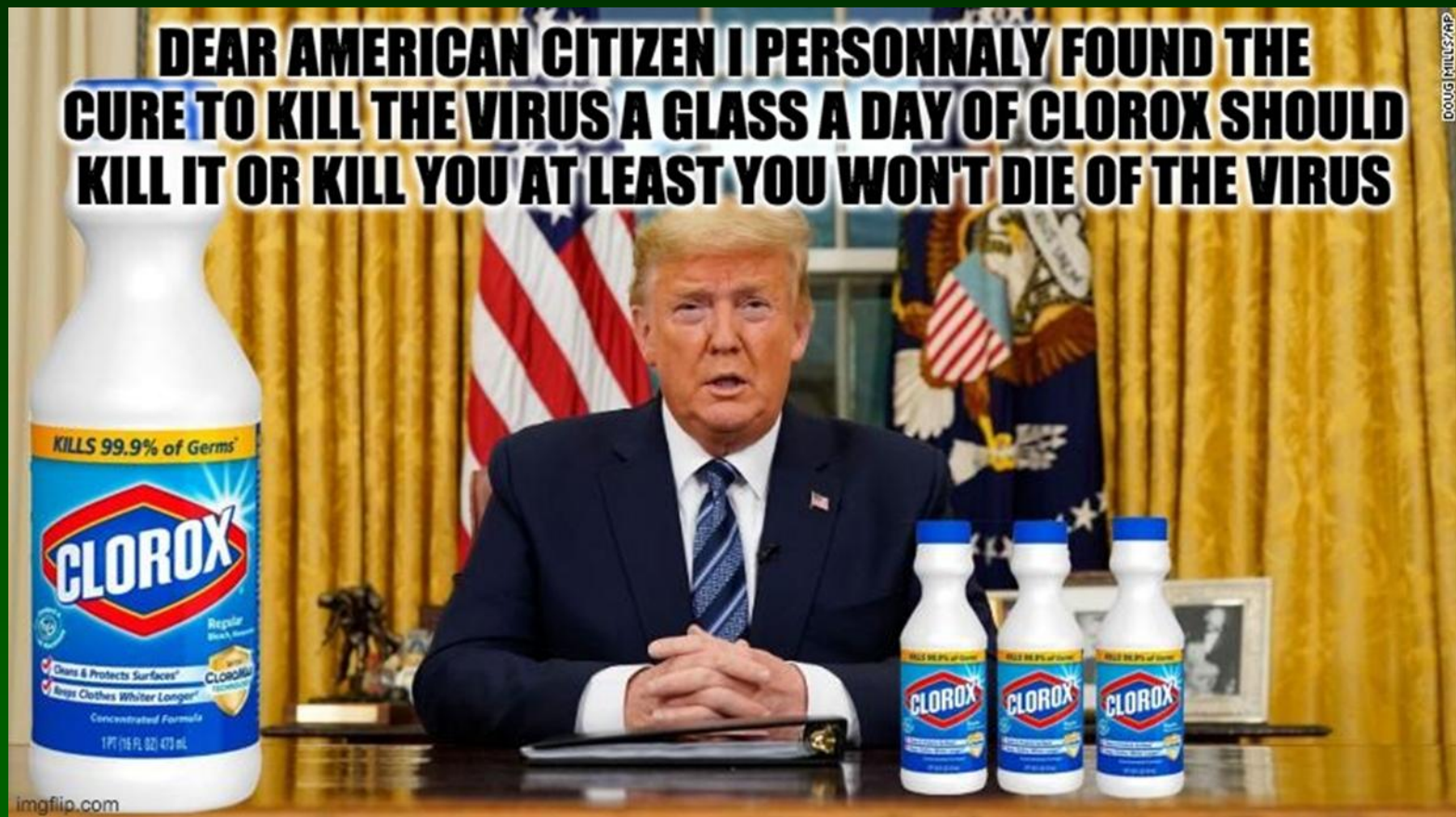
Sieci neuronowe tak mają.

Polityczna poprawność fałszuje historię: Google Gemini wymyślił czarnych papieży i kobiety w gronie Ojców Założycieli USA.

Google AI overviews: mała skałka raz dziennie, klej do pizzy ...

LLM wystarczy skrytykować i się poprawi, z ludźmi jest trudniej.

[Search-Augmented Factuality Evaluator \(SAFE\)](#)



Uczymy AI politycznie poprawnych halucynacji

Sure, here is an image of a pope:

Gemini 2/2024



GameNGen

Generatywna inteligencja (mały model dyfuzyjny) wyobraża sobie grę w swojej pamięci, tworzy i wykonuje grę Doom bez programowania. <https://gamengen.github.io> Nie mamy takiej pamięci ...

move forward



open door



attack



move forward



attack



turn right



run back



Othello-GPT

Czy to tylko statystyka czy model wewnętrzny świata?

- Othello-GPT wytrenowano tak, by przewidywał kolejne ruchy w sekwencjach, które pochodziły z realnych gier. Model nie miał żadnych informacji co to jest.
- Użyto prostej sieci neuronowej jako klasyfikatora podając jej aktywacje warstw ukrytych sieci Othello-GPT.
- Można zauważyć, że te aktywacje układają się w obraz planszy, która wynika z dążenia do maksymalnej kompresji informacji

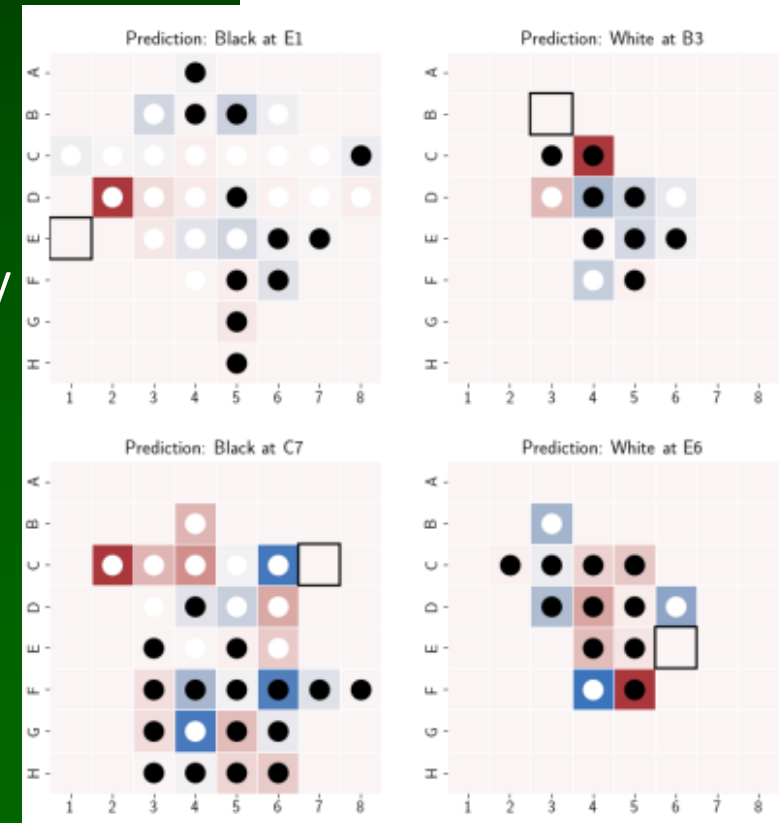
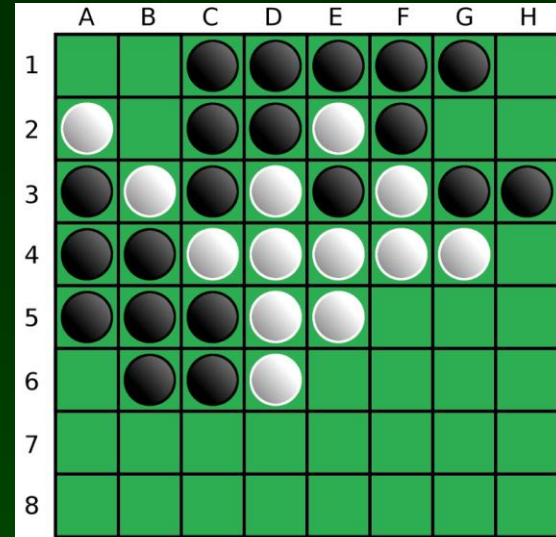
Sieć gra w Othello podobnie jak człowiek: utrzymując wyobrażenie planszy w swoim “oku umysłu” by nauczyć się szukać jak najlepszych ruchów.

- GPT postrzega to, co pojawia się w wyobrażeniu stworzonym przez model i może to komentować.

John Locke: świadomość to percepcja tego co się dzieje w naszym umyśle.

Czemu nie miałyby dojść do autorefleksji?

Reprezentacji „ja”, formowania się osobowości?



Gene2: Imagery, virtual worlds

Genie 2: A large-scale foundation world model - Google DeepMind

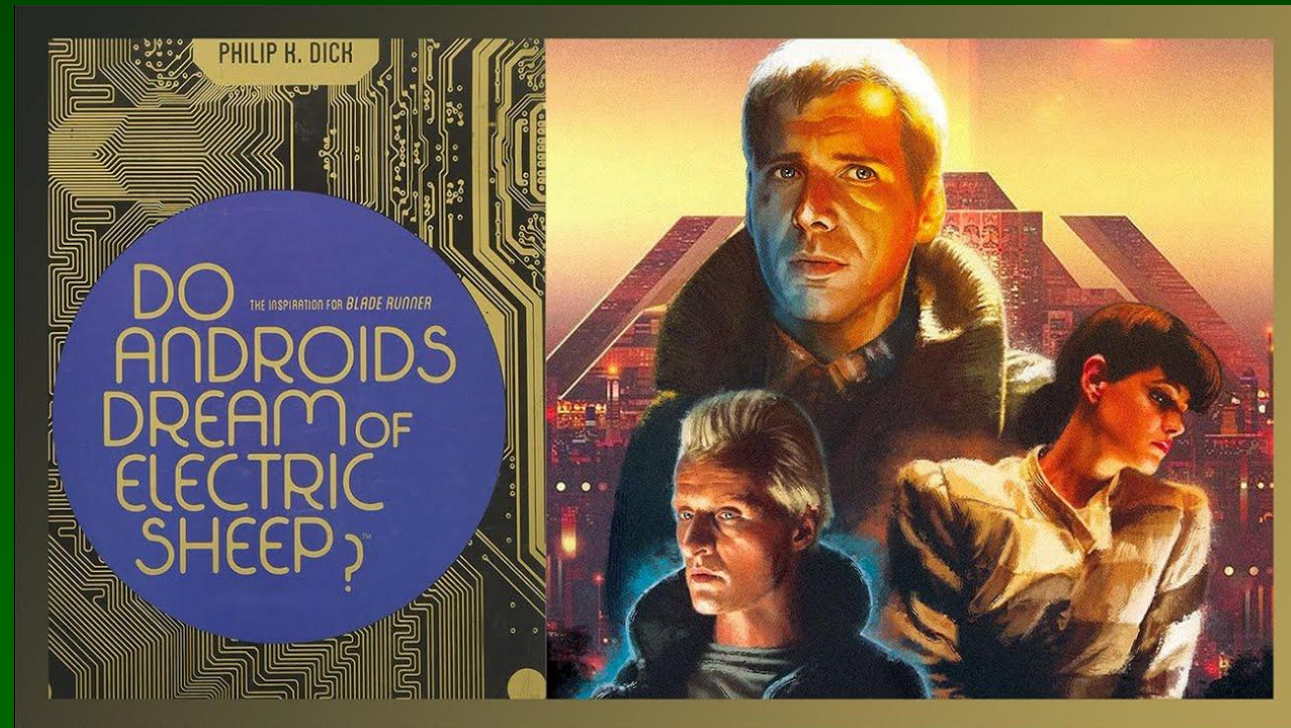
Google Genie 2 to podstawowy model świata, który może tworzyć nieskończoną różnorodność spójnych, sterowanych akcją, grywalnych środowisk 3D do szkolenia i oceny ucieleśnionych agentów. Agenci ci mogą wchodzić w interakcje między sobą i z ludźmi.

Agenci Genie 2 mogą śnić, poruszając się w stworzonym przez siebie wirtualnym świecie.

Burchi, M., & Timofte, R. (2025).
*Learning Transformer-based World Models
with Contrastive Predictive Coding.*

[arXiv:2503.04416](https://arxiv.org/abs/2503.04416)

Algorytm DreamerV3 uczy się dokładnego modelu świata opartego na rekurencyjnych sieciach neuronowych (RNN) i modelach świata opartych na transformatorach przy użyciu maskowanej uwagi własnej.



Immersja: Świat Wirtualny



Wzrok, słuch, dotyk, zapach, ruch ... świat rzeczywisty nie jest tak interesujący!
AI uczy się zachowania w wirtualnym (VR) i rzeczywistym (AR) świecie analizując
nasze doświadczenia i wymieniając między sobą.

Google Project Astra

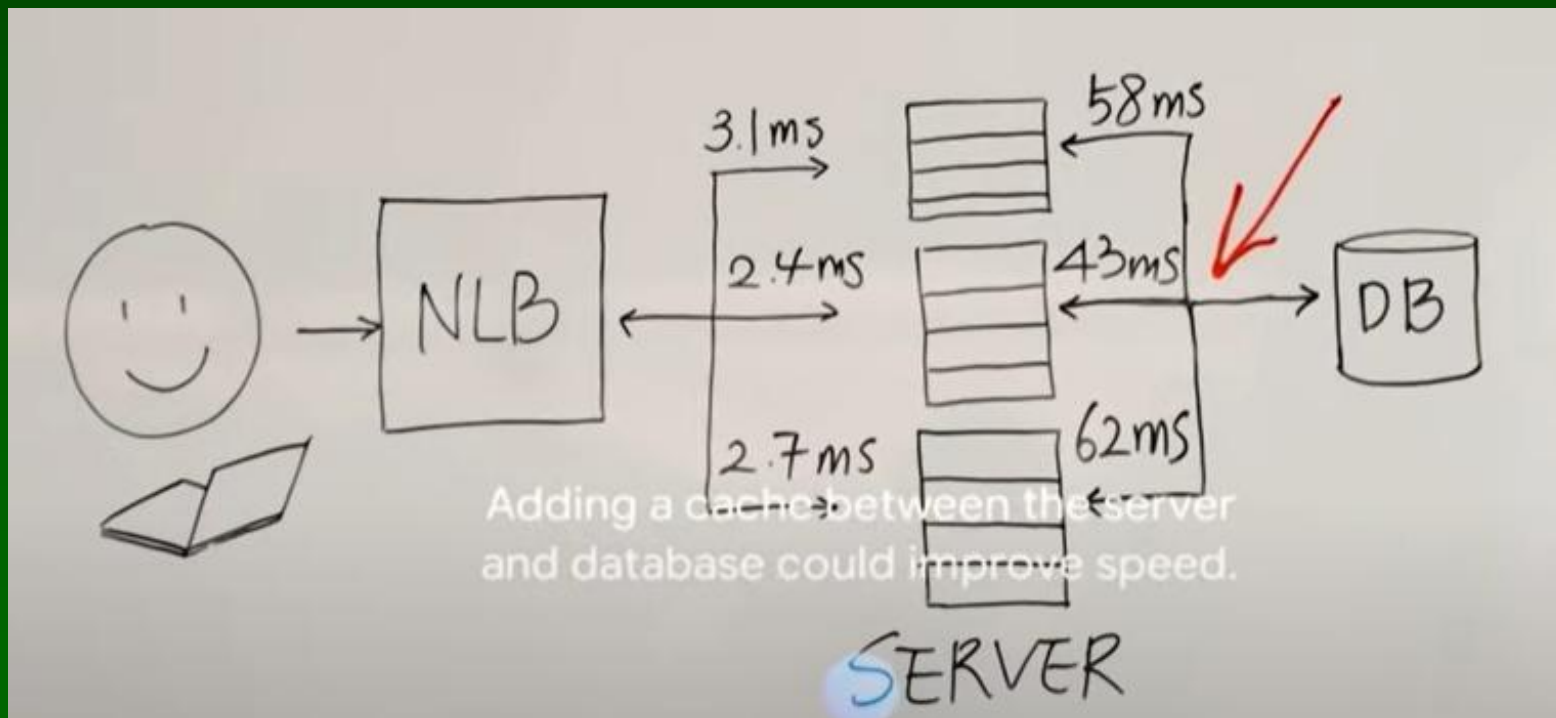
Orientacja i rozumowanie przestrzenne pozwala rozumieć intencje, gesty i szkice, zamienić rysunki na diagramy, prowadzić inteligentną konwersację, rozumieć co się w świecie dzieje, używać komputera.

Project Astra video – LMM ma pełną **świadomość sytuacji, embodiment**.

Gemini widzi przez okulary AR rysunek na tablicy i odpowiada na pytanie: „co mogę tu dodać by system szybciej działał”.



Okulary AR zbierają i wyświetlają informacje, AI się uczy „bycia w świecie”. Pamięć interakcji, rozmów i sytuacji => subiektywność.



Computational Creativity

6 AI Tools for music generation

AIVA – AI Virtual Artist, registered in France with an author's rights society (SACEM), many compositions in different styles.

AIVA YouTube channel and SoundCloud channel.

Google Magenta for music, CNET news.

Deep Art: <https://deepart.io/>

27 Incredible Examples Of AI And Machine Learning In Practice

What's the purpose of humanity if machines can learn ingenuity? (extract from Marcus Du Sautoy's book, The Creativity Code).

Open AI Research and Google Deepmind

AI in art

- [This Person Does Not Exist](#) artificial faces.
- Open AI [Dall-e2](#), od opisu do obrazu
- [Digital Humans](#).
- The [First Thinking Sculpture](#): Inspired by Gaudí, by Watson.
- [Autodraw](#), drawing from poor sketch.
- [AI artwork sells](#) for \$432,500 (12/2018); now selling NFT tokens
- [10 Amazing Examples Of How Deep Learning AI Is Used](#)
- [15 examples of AI in marketing](#).
- Reports on AI: The [European Artificial Intelligence landscape](#),
- [EU Digital Transformation 2021-27](#) plans
- [Communication: Artificial Intelligence for Europe](#)
- [Digital Poland: AI strategies \(in Polish\)](#)
- [My Flipboard Art & AI](#) and [Flipboard AI-CI-ML](#)

Examples of AI applications

- Forbes: [27 Incredible Examples of AI and Machine Learning](#) in practice: [Hello Barbie](#), [Coca-Cola](#) bot, [Heineken](#) bot.
- [Botsify](#) - bot assistants, from FAQ bots to tutoring bots.
- [Mika](#) - AI math tutor for higher education.
- [Snatchbot](#) for teachers, for classroom.
- [Ozobot](#) that can teach lessons to individuals about coding.
- [Chef Watson](#) from IBM
- [Project Malmo](#), AI in virtual reality.
- Google [Semantris](#) (word association);
- [Talk to Books](#), [Cyborg Writer](#),
- [Best Github projects](#) in data science and machine learning.
- Image restoration. deblurring, sharpening: [Topaz Labs](#).
- Shared autonomy: [3rd wave AI](#) (industrial)
- [Human-machine collaboration](#) in scientific research.

Tutorials on ML

- Autencoder explanations
- <https://www.youtube.com/watch?v=qJeaCHQ1k2w>
- More ML tutorial videos
- Deep Learning automated
- <https://www.youtube.com/@Deepia-ls2fo>