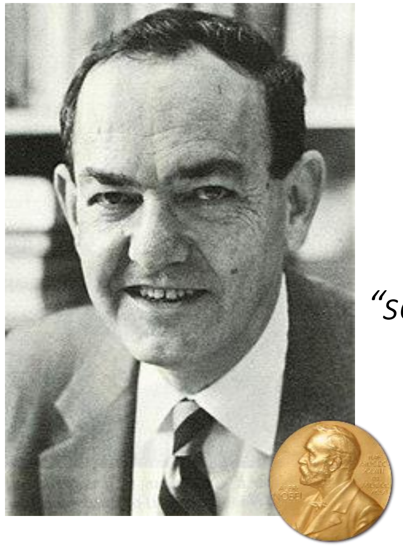
Tutorial Outline

- **Part 1:** Background and challenges (20 min)
- **Part 2:** Preliminaries of invariance (20 min)
- *Q&A / Break (10 min)*
- **Part 3:** Invariance in the era before deep learning (30 min)
- **Part 4:** Invariance in the early era of deep learning (10 min)
- *Q&A / Coffee Break (30 min)*
- **Part 5:** Invariance in the era of rethinking deep learning (50 min)
- **Part 6:** Conclusions and discussions (20 min)
- *Q&A (10 min)*

A Historical Perspective of Data Representation
Rethinking Deep Learning with Invariance: The Good, The Bad, and The Ugly

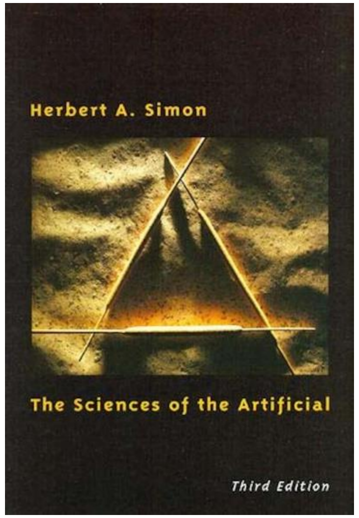
Deep (Representation) Learning,
A Big Bang Moment For AI

Data Representation



H. Simon, 1969
The Sciences of the Artificial

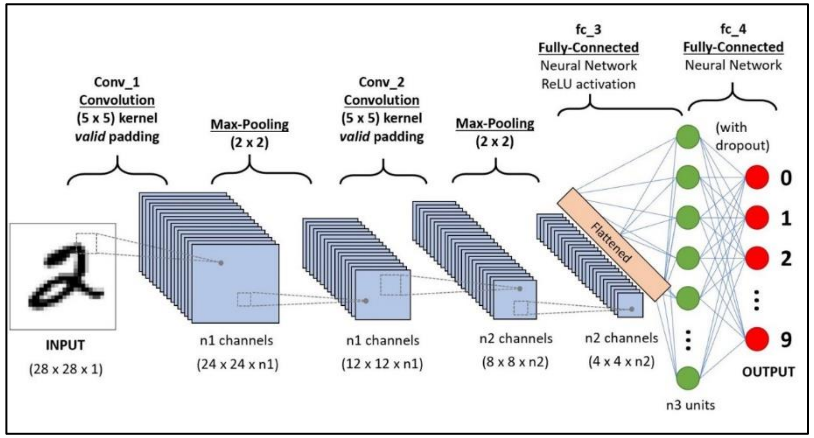
“solving a problem simply means representing it so as to make the solution transparent”

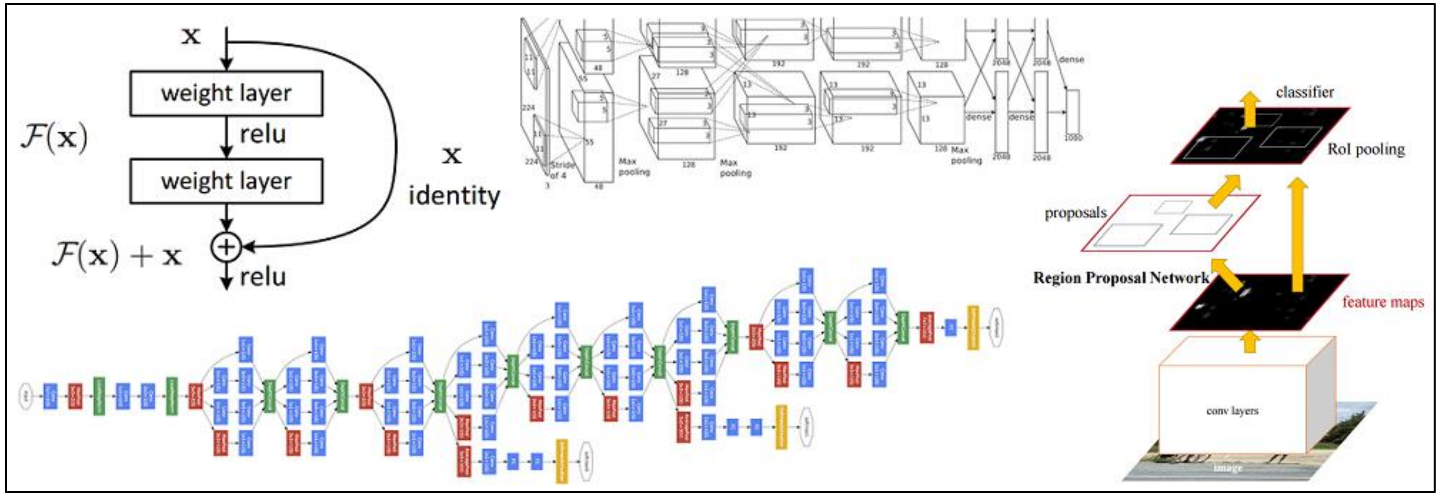


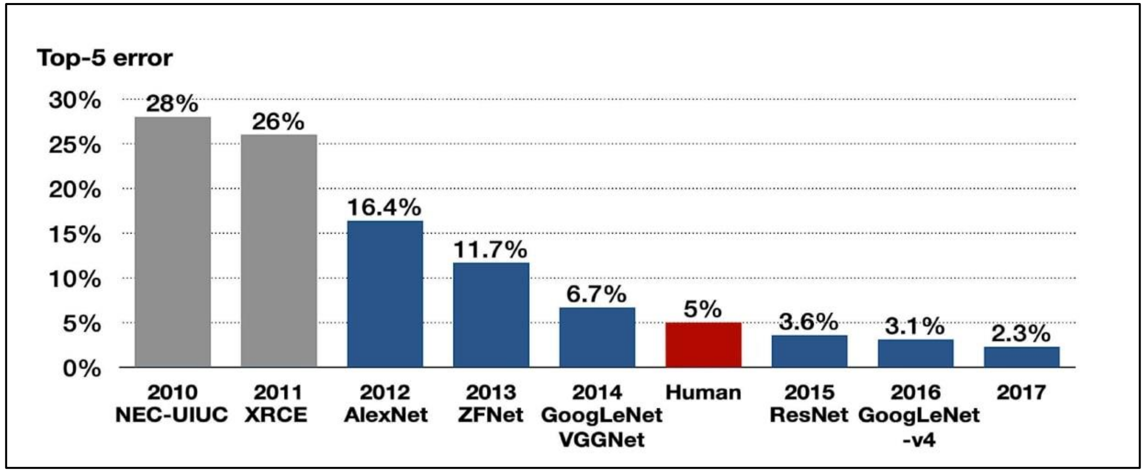
Data ⇨ Representation ⇨ Knowledge discovery ⇨ Application

- H Simon. *The Sciences of the Artificial (Third edition)*. MIT Press, 1996.

Processing Human Perceptual Information



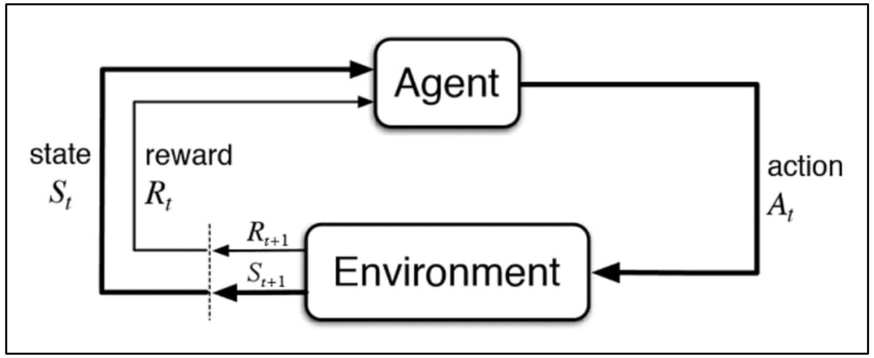


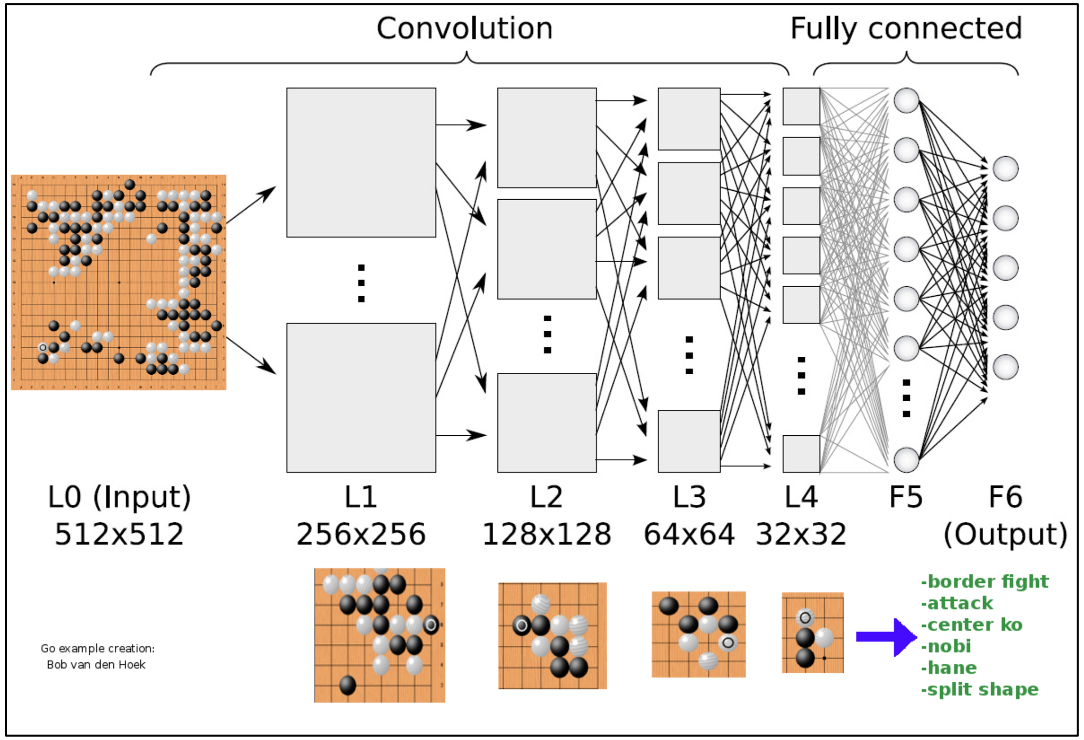


- J Deng, W Dong, R Socher, et al. ImageNet: A large-scale hierarchical image database. *CVPR,* 2009.

Playing Board Games

AlphaGo





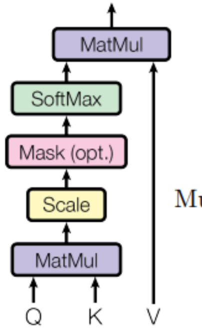
- D Silver, J Schrittwieser, K Simonyan, et al. Mastering the game of go without human knowledge. *Nature,* 2017.

Generating Realistic Media

ChatGPT

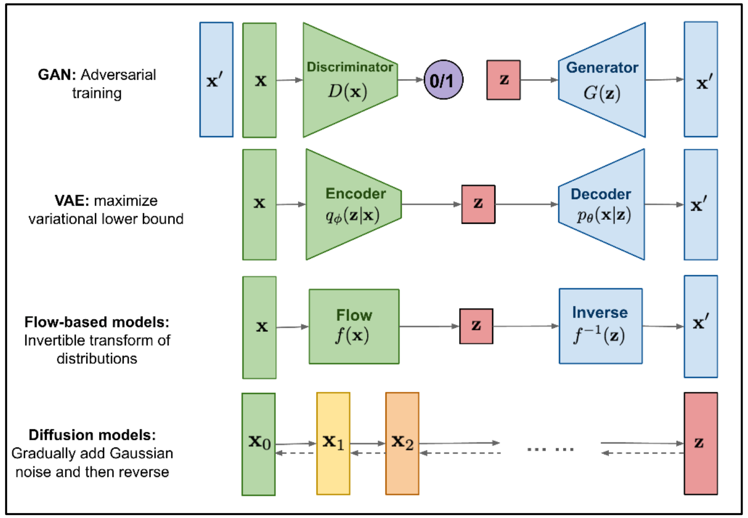
DALL·E 3

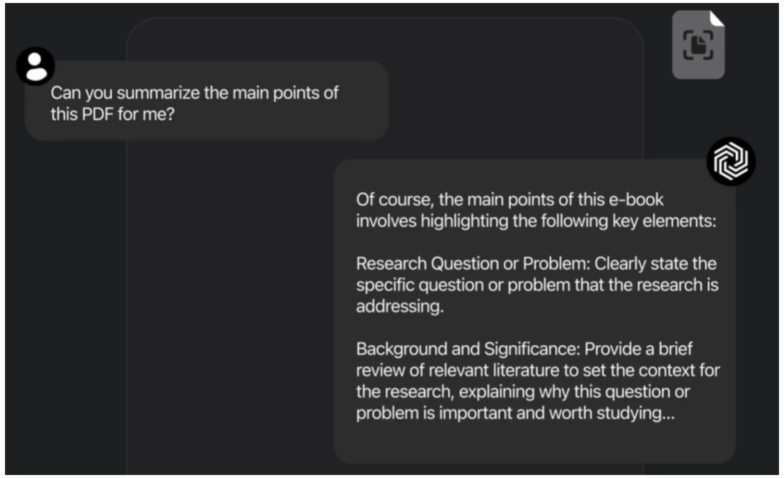
OpenAI SORA



$$\text{Attention}(Q, K, V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V$$

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, ..., \text{head}_\text{h})W^O$$
$$\text{where head}_\text{i} = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$$

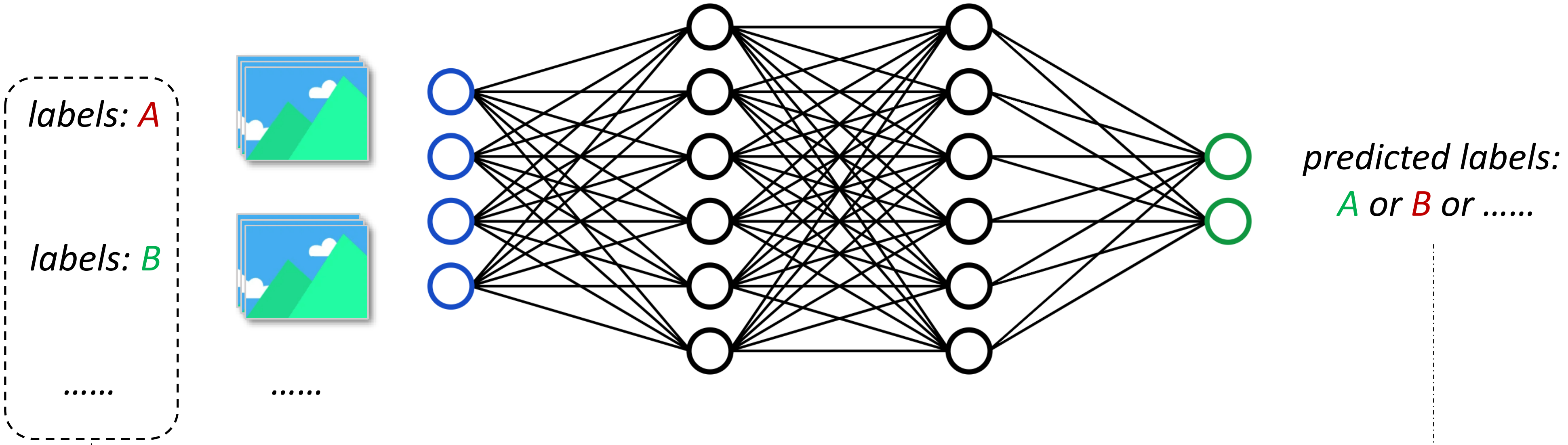




- Z Epstein, A Hertzmann, L. Herman, et al. Art and the science of generative AI. *Science*, 2023.

Empirical Risk Minimization (ERM), Behind All These Successes

Empirical Risk Minimization



$$\mathbb{E}_{\mathbf{x},\mathbf{y}\in S}[\ell(\mathbf{y}, h(\mathbf{x}))]$$

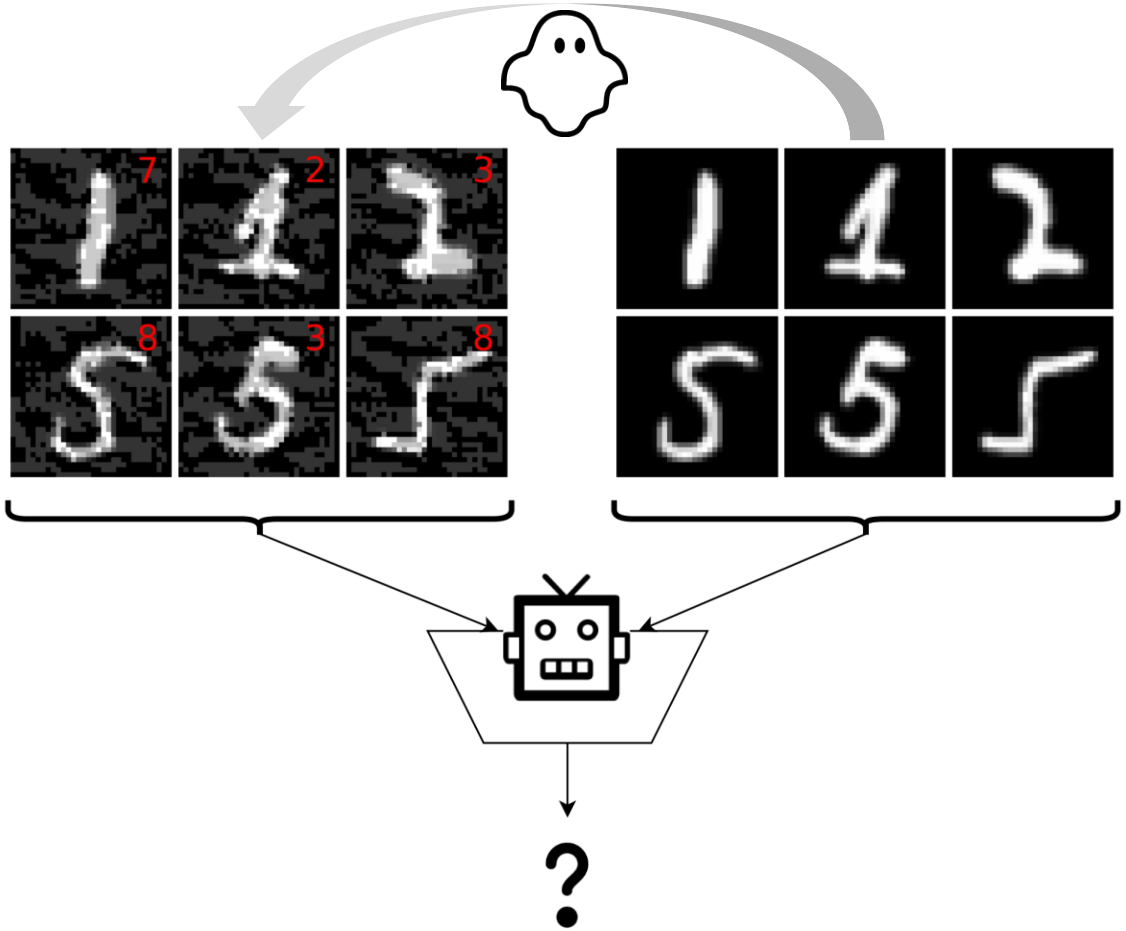
Empirical learning draws the lines between categories.

But what about robustness, interpretability, and efficiency?

- V Vapnik. Principles of risk minimization for learning theory. *NIPS,* 1991.

Robustness of Empirical Learning

- Robustness: the performance of a system is stable for intra-class variations on the input.



Universal

Color

One-pixel

Watermark

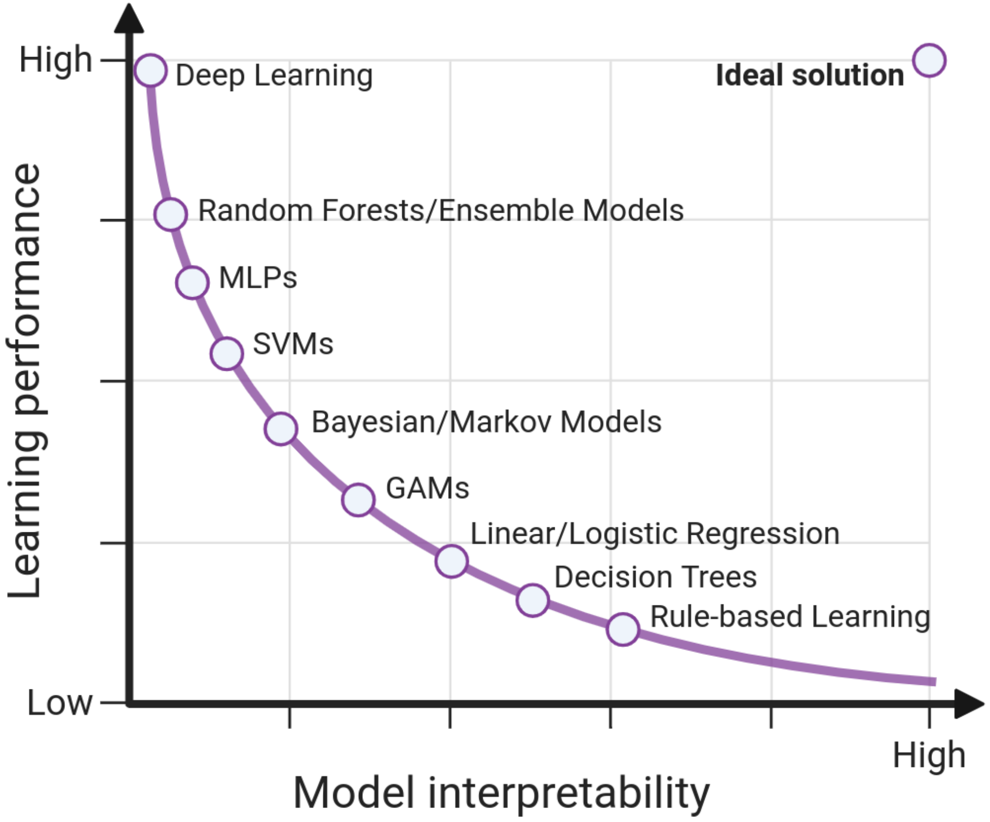
Physical

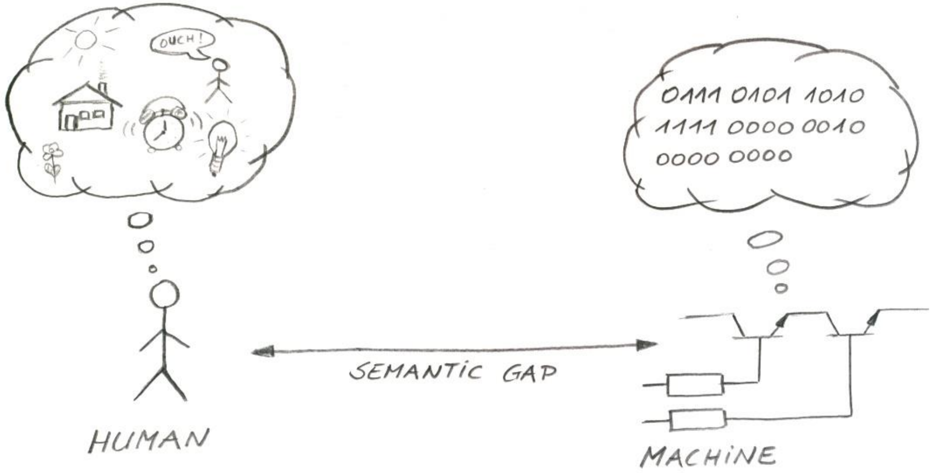
Weird

- C Buckner. Understanding adversarial examples requires a theory of artefacts for deep learning. *Nature Machine Intelligence*, 2020.

Interpretability of Empirical Learning

- Interpretability: the behavior of a system can be understood or predicted by humans.

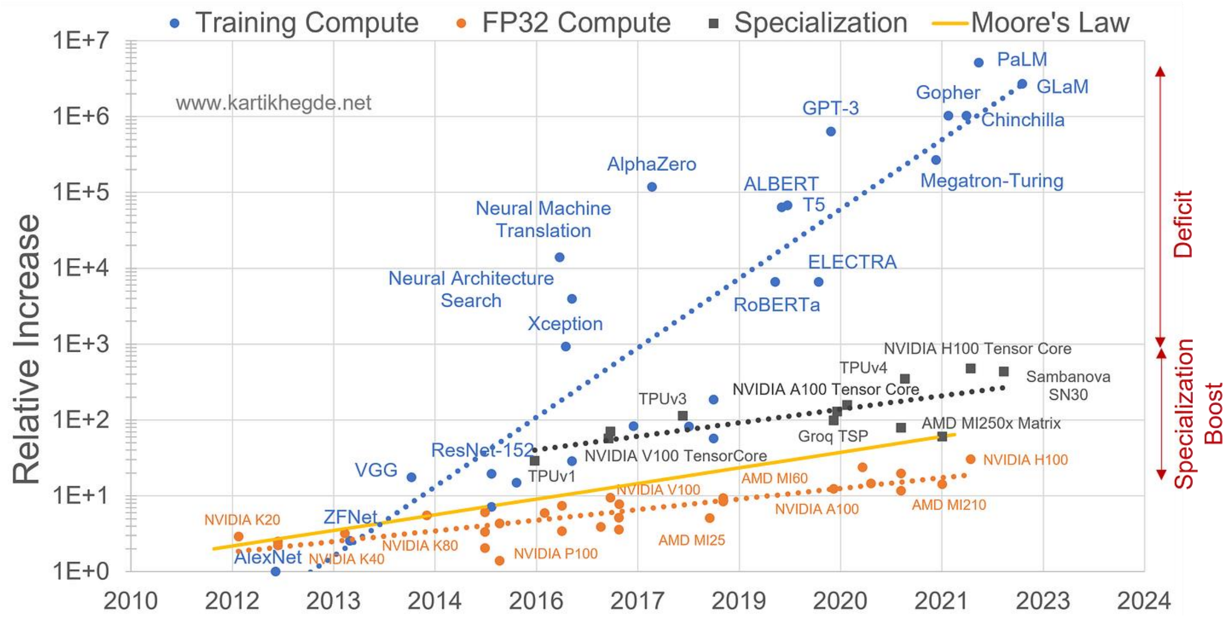


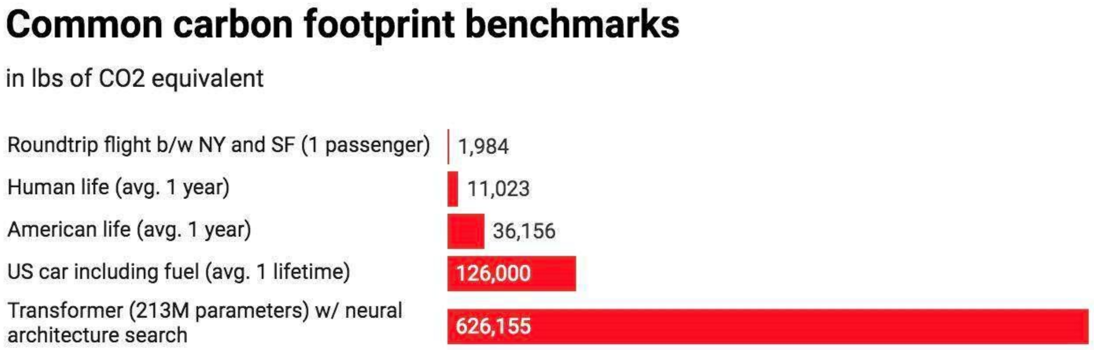


- X Li, C Cao, Y Shi, et al. A survey of data-driven and knowledge-aware explainable AI. *TKDE,* 2020.

Efficiency of Empirical Learning

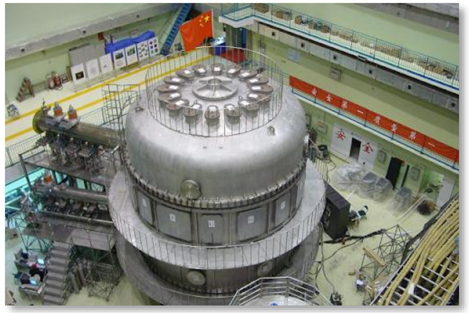
- Efficiency: the real-time availability and energy cost during human-computer interaction.



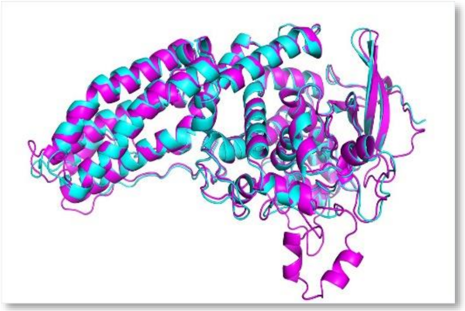


- E Strubell, A Ganesh, A McCallum, et al. Energy and policy considerations for modern deep learning research. *AAAI,* 2020.

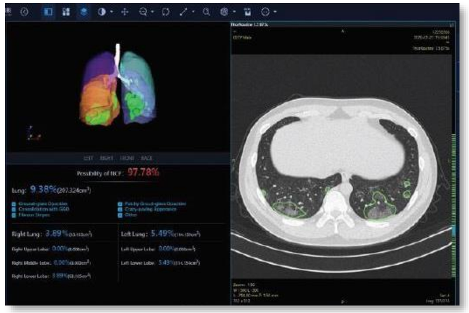
When Moving Towards Trustworthy AI



Tokamak Control



Drug Discovery



Imaging Diagnostics

Automatic Driving

Cyber Security

Biometrics

Empirical learning *v.s.* robustness, interpretability, efficiency…

- H Liu, M Chaudhary, H Wang. Towards trustworthy and aligned machine learning: A data-centric survey with causality perspectives. *arXiv preprint arXiv:2307.16851,* 2023.

A Fundamental Prior Underlying Both Natural World And AI Systems

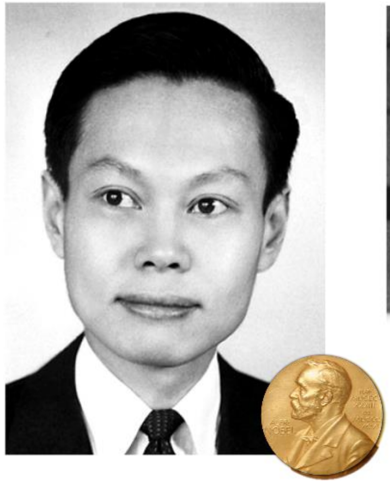
Invariance/Symmetry in Natural World

- A symmetry of a system is a transformation that leaves a certain property invariant.

F. Klein, 1872
Erlangen Program

E. Noether, 1918
Noether's Theorem

H. Weyl, 1929
The Book of Symmetry



C. N. Yang & R. L. Mills, 1954
Yang-Mills Theory

- F Klein. A comparative review of recent researches in geometry. *Bulletin of the American Mathematical Society*, 1893.
- H Weyl. *Symmetry.* Princeton University Press, 2015.

Invariance/Symmetry in AI Systems

- An AI system is a digital modeling of the physical systems in the natural world.

Y. LeCun, Y. Bengio & G. Hinton, 2015,
Deep learning, Nature

The Selectivity–Invariance Dilemma:
“representations that are selective to the aspects that are important for discrimination, but that are invariant to irrelevant aspects”

- Y LeCun, Y Bengio, G Hinton. Deep learning. *Nature*, 2015.
- Y Bengio, A Courville, P Vincent. Representation learning: A review and new perspectives. *TPAMI*, 2013.

How Invariance/Symmetry Helps Robustness, Interpretability, Efficiency

- Perfect robustness — the performance of the AI system remains invariant with respect to the transformations of interest.
- Interpretable concept — humans and AI systems share a basic concept that allows humans to predict AI behavior on transformations of interest.
- Structural efficiency – AI systems no longer need to memorize non-discriminative data variants.