

# The Second Curve of Scaling Law

[aka.ms/GeneralAI](https://aka.ms/GeneralAI)

Furu Wei, Partner Research Manager

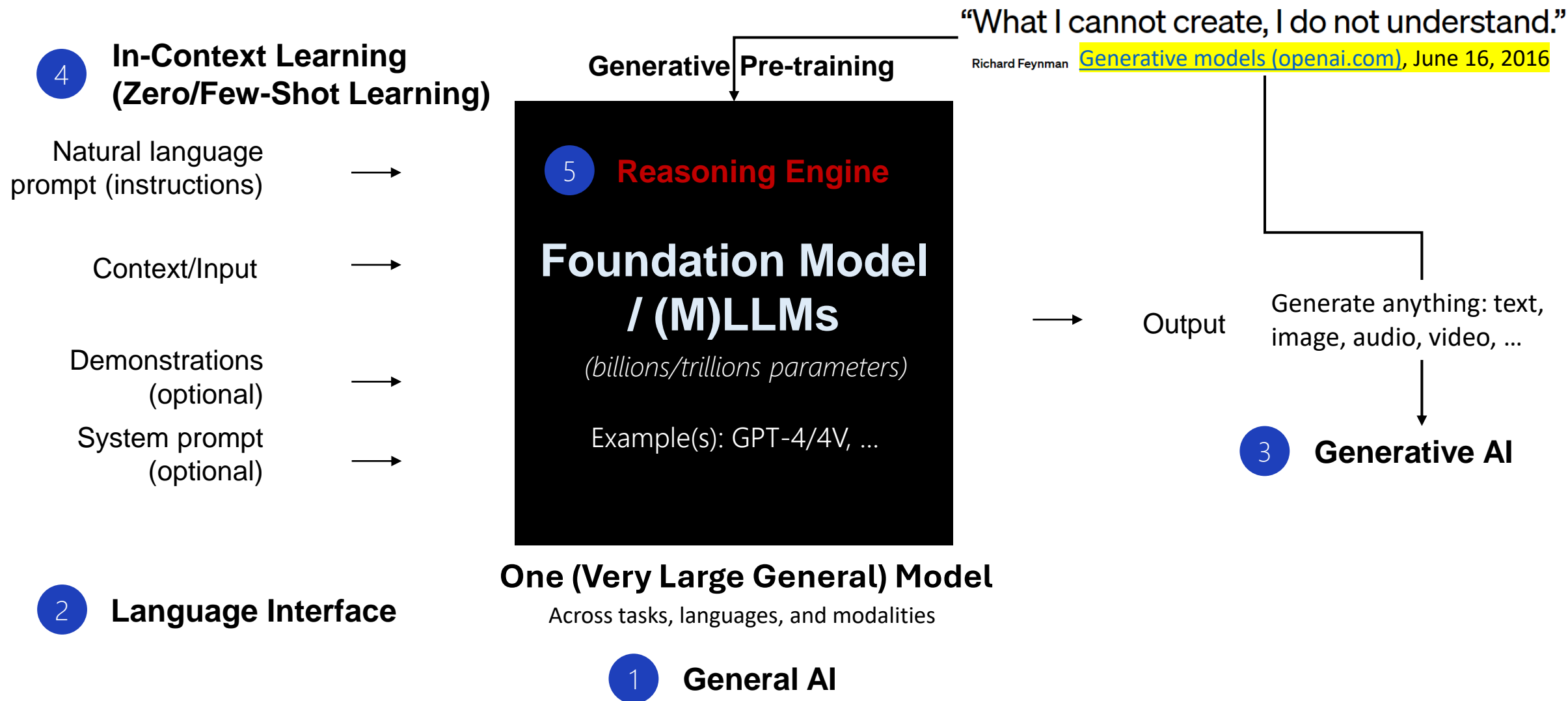
Microsoft Research

Dec. 2023

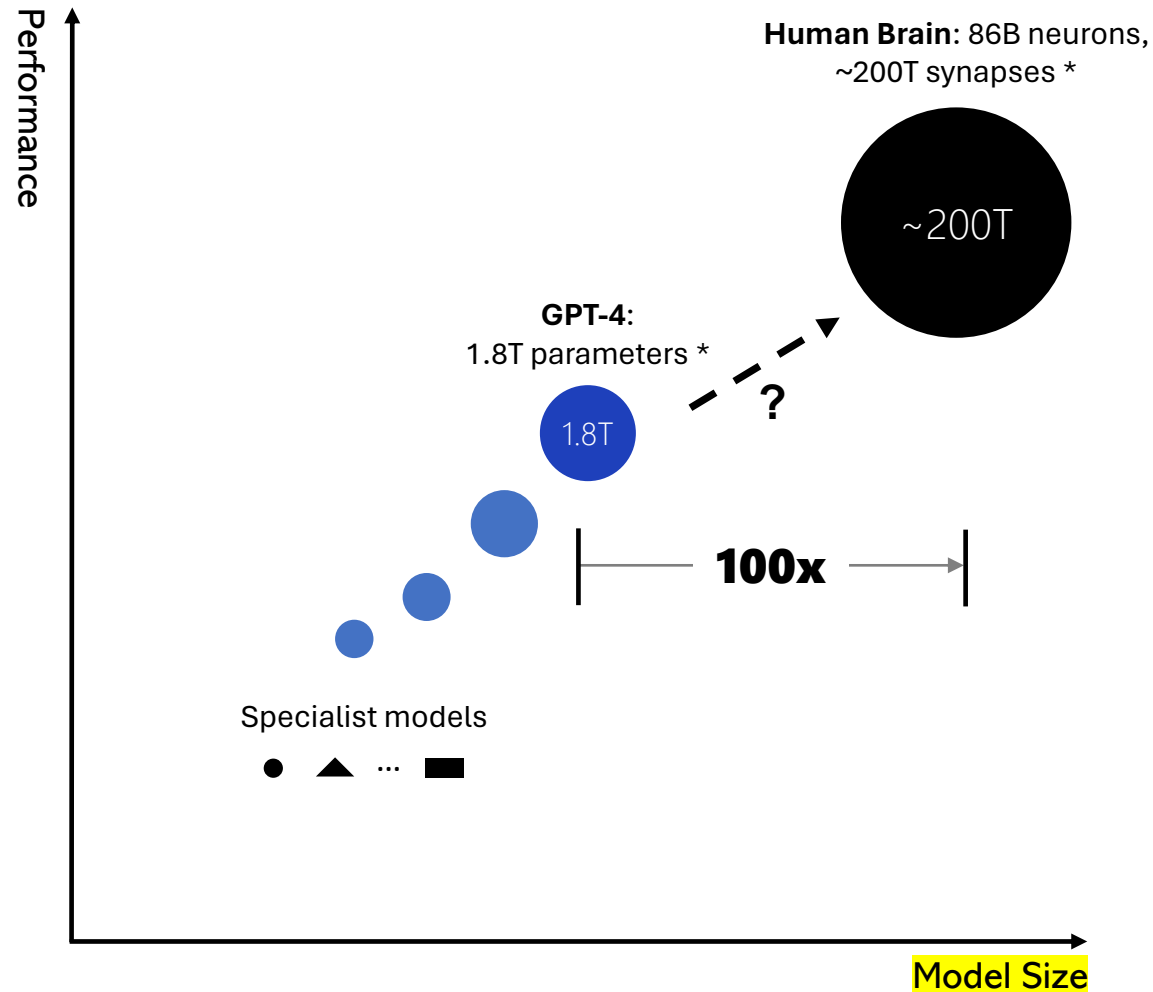
# The New Age of AI (2020/GPT-3 - )

Paradigm Shift

2022/ChatGPT



# The Power of Scaling (Law)



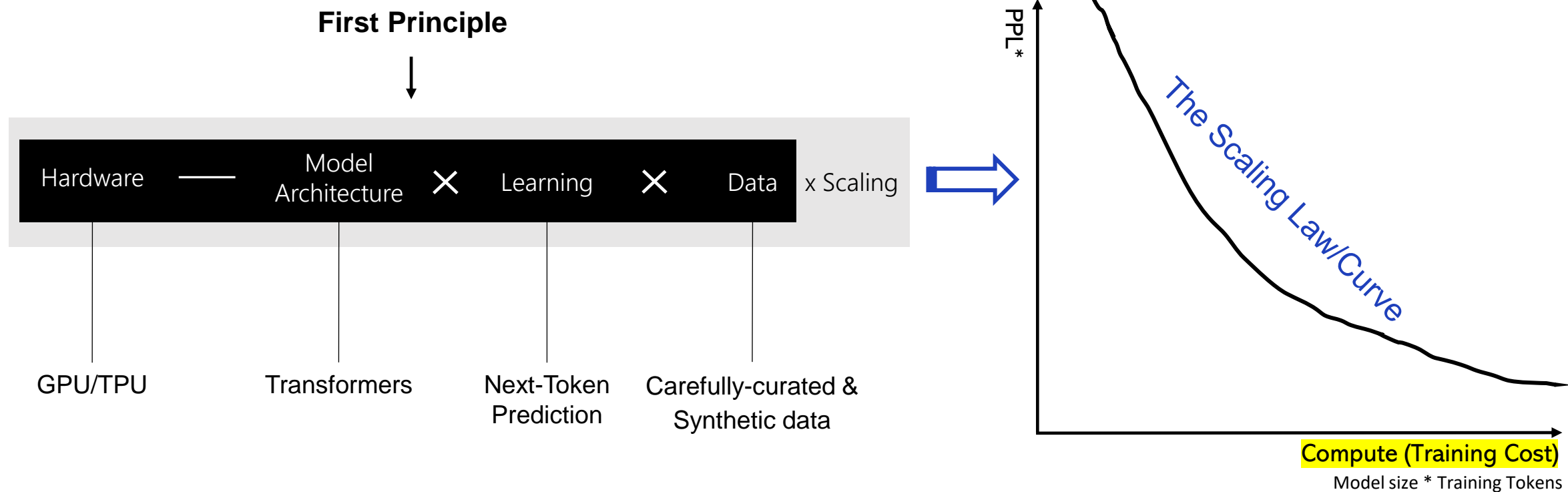
Is (the first) scaling (curve) all we need?

**YES** and **NO**

\* [GPT-4 architecture, datasets, costs and more leaked](#)

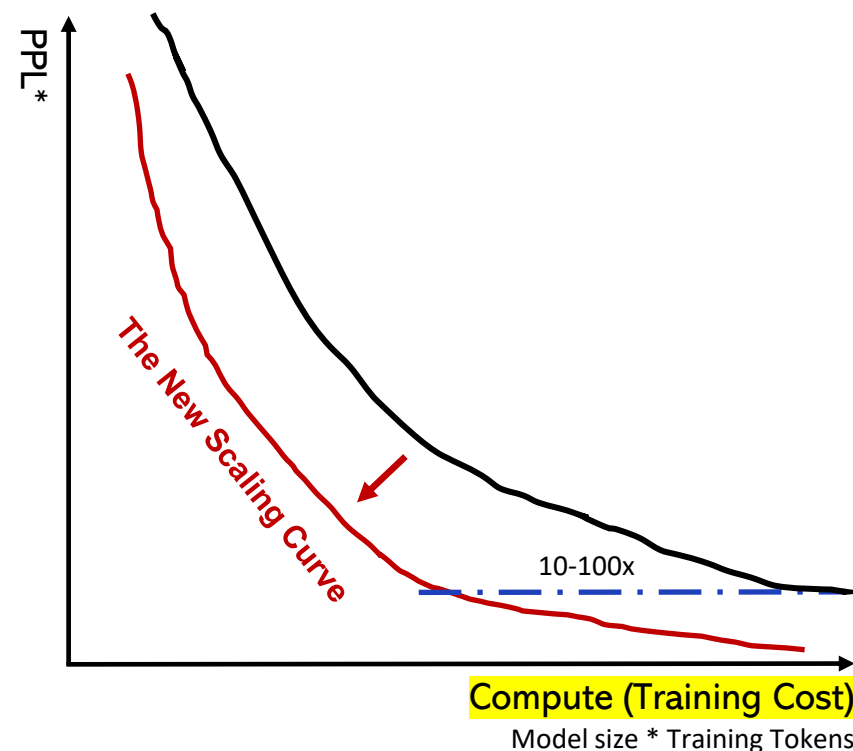
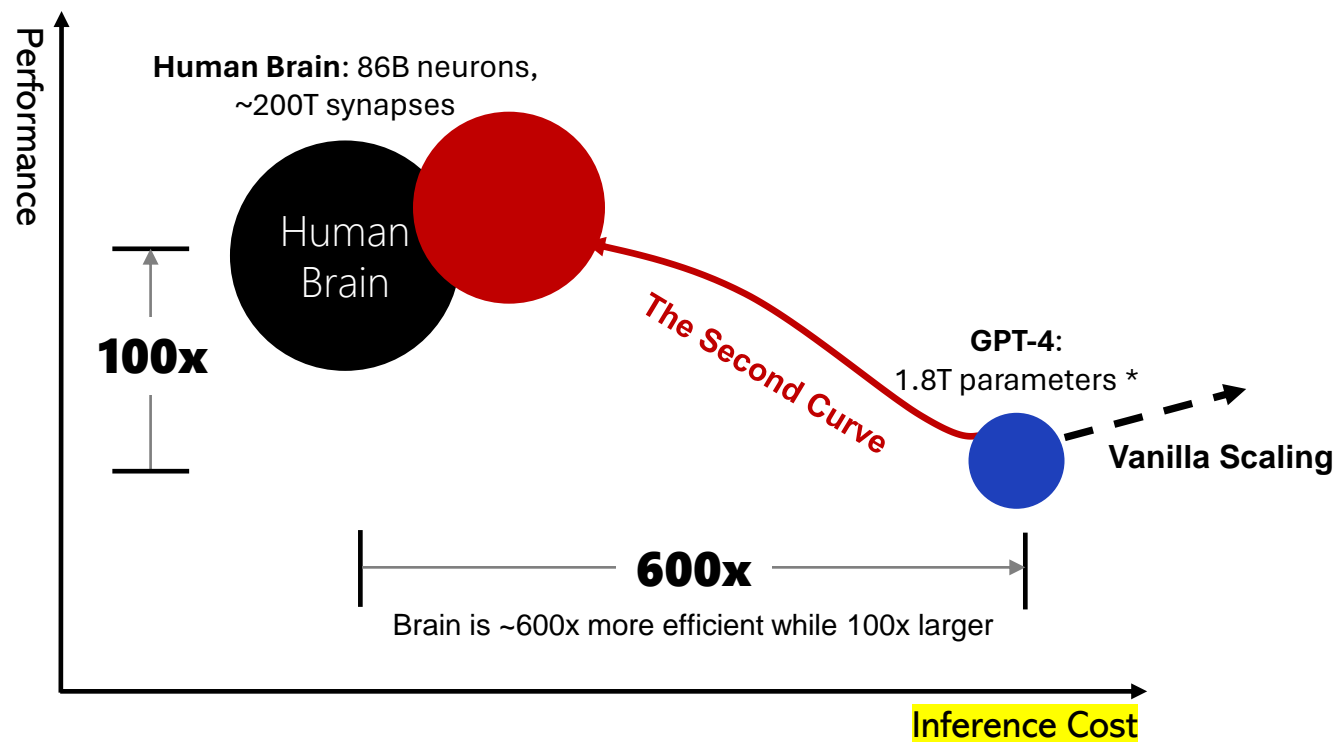
\* [List of animals by number of neurons - Wikipedia](#)

# First Principle of Scaling Law (Curve)



\* Lower PPL = higher compression ratio / performance / intelligence

# The Second Curve of Scaling Law



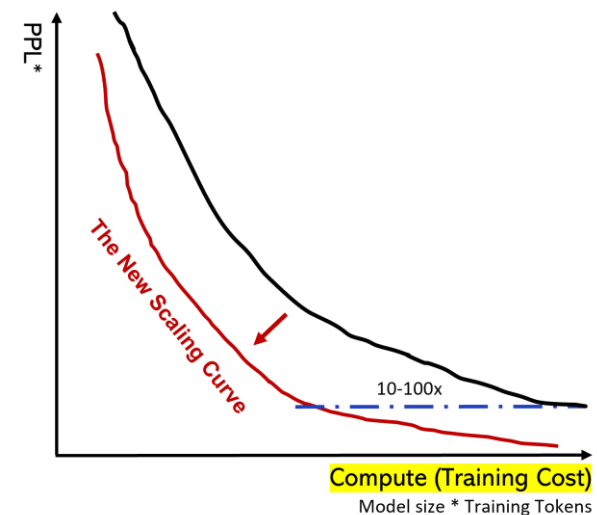
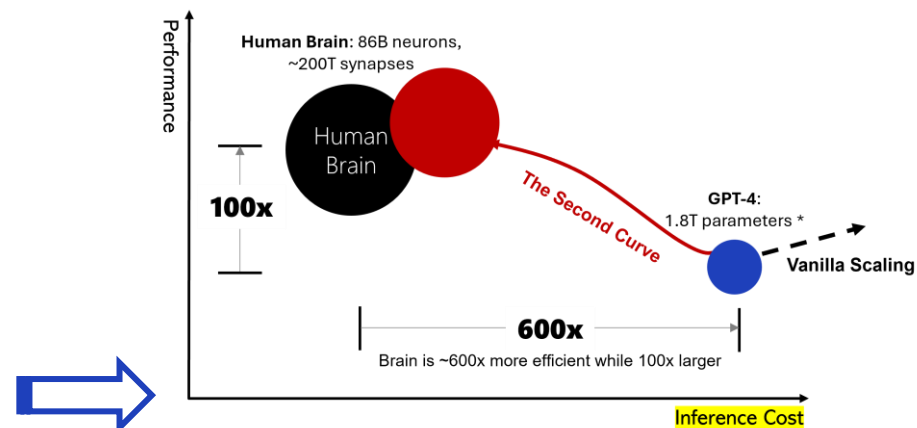
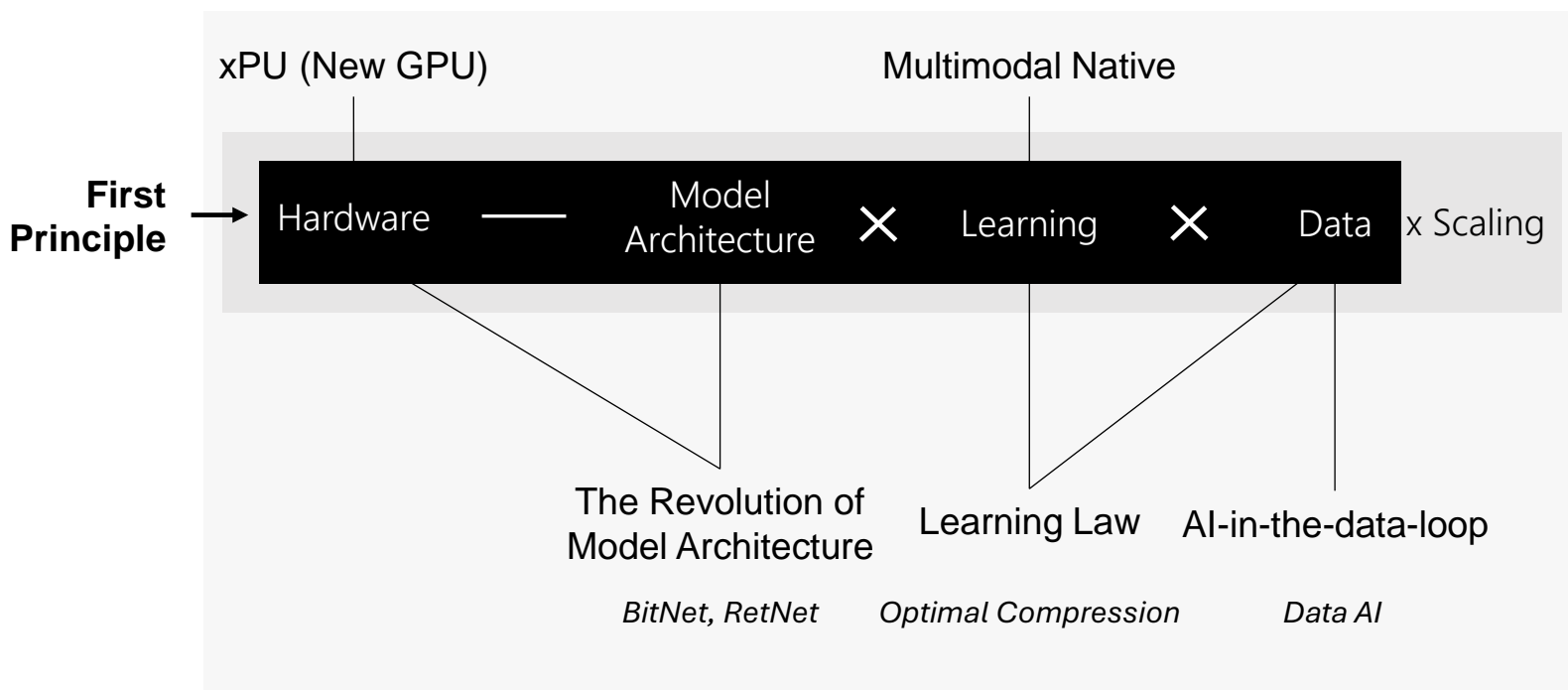
\* [GPT-4 architecture, datasets, costs and more leaked](#)

\* [List of animals by number of neurons - Wikipedia](#)

\* [Does Thinking Really Hard Burn More Calories? - Scientific American](#)

\* Lower PPL = higher compression ratio / performance / intelligence

# The Second Curve of Scaling Law



aka.ms/GeneralAI

Contact: Furu Wei (fuwei AT microsoft.com)