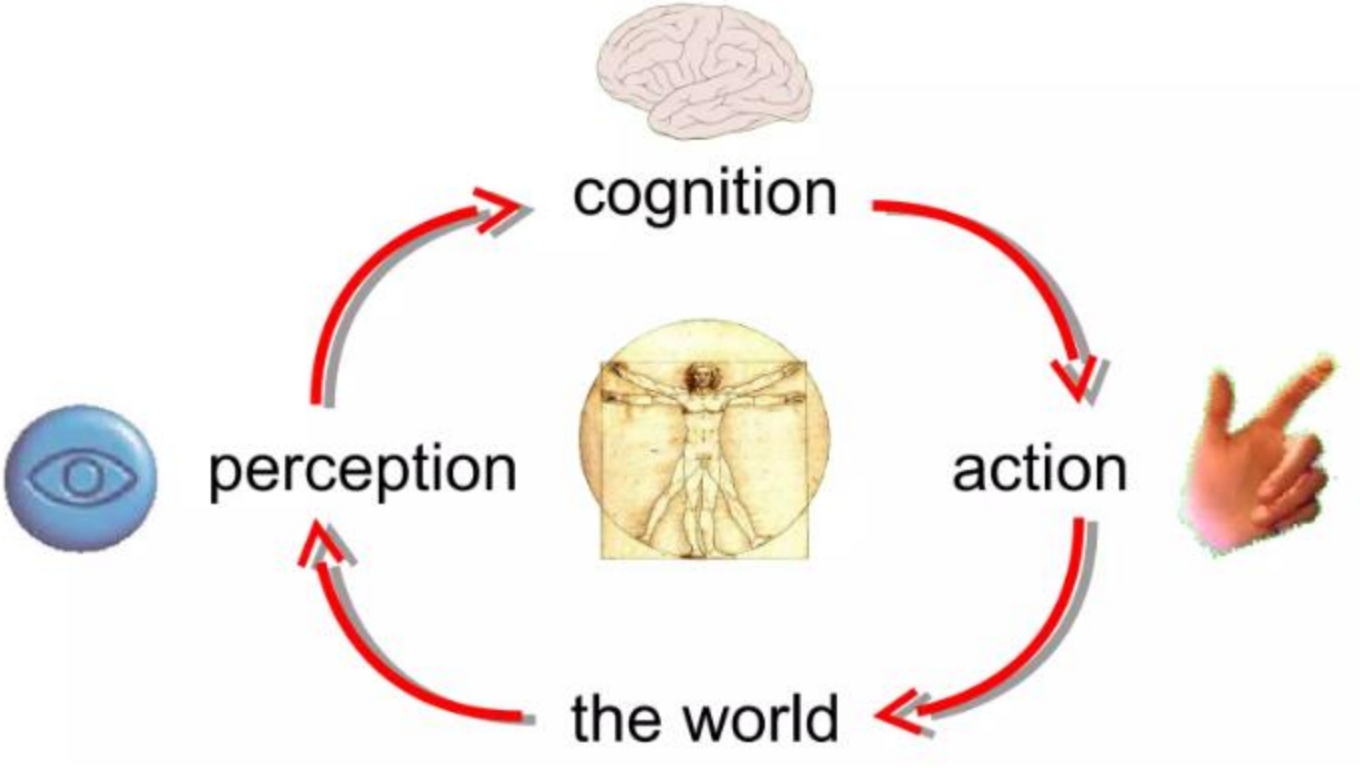
Artificial Intelligence
(Big-picture)

Human Agent

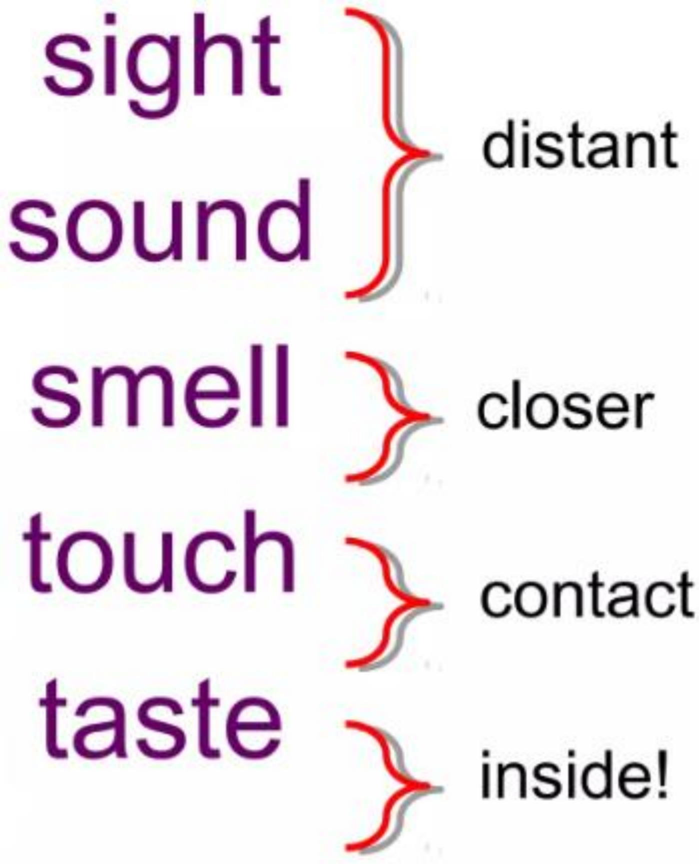
we live to act



Human Agent

- Perception
- Cognition
  - Memory
  - Thinking
    - Learning -> Knowledge
    - Reasoning -> Decision Making
    - Imagination
- Action
- Autonomous

Perception



sight

sound

distant

smell

closer

touch

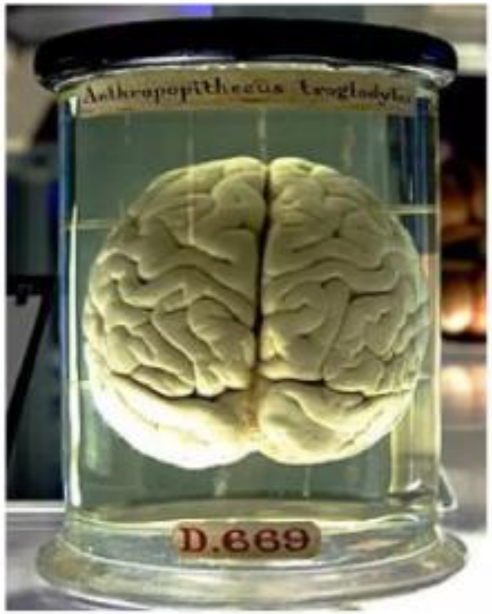
contact

taste

inside!

Cognition

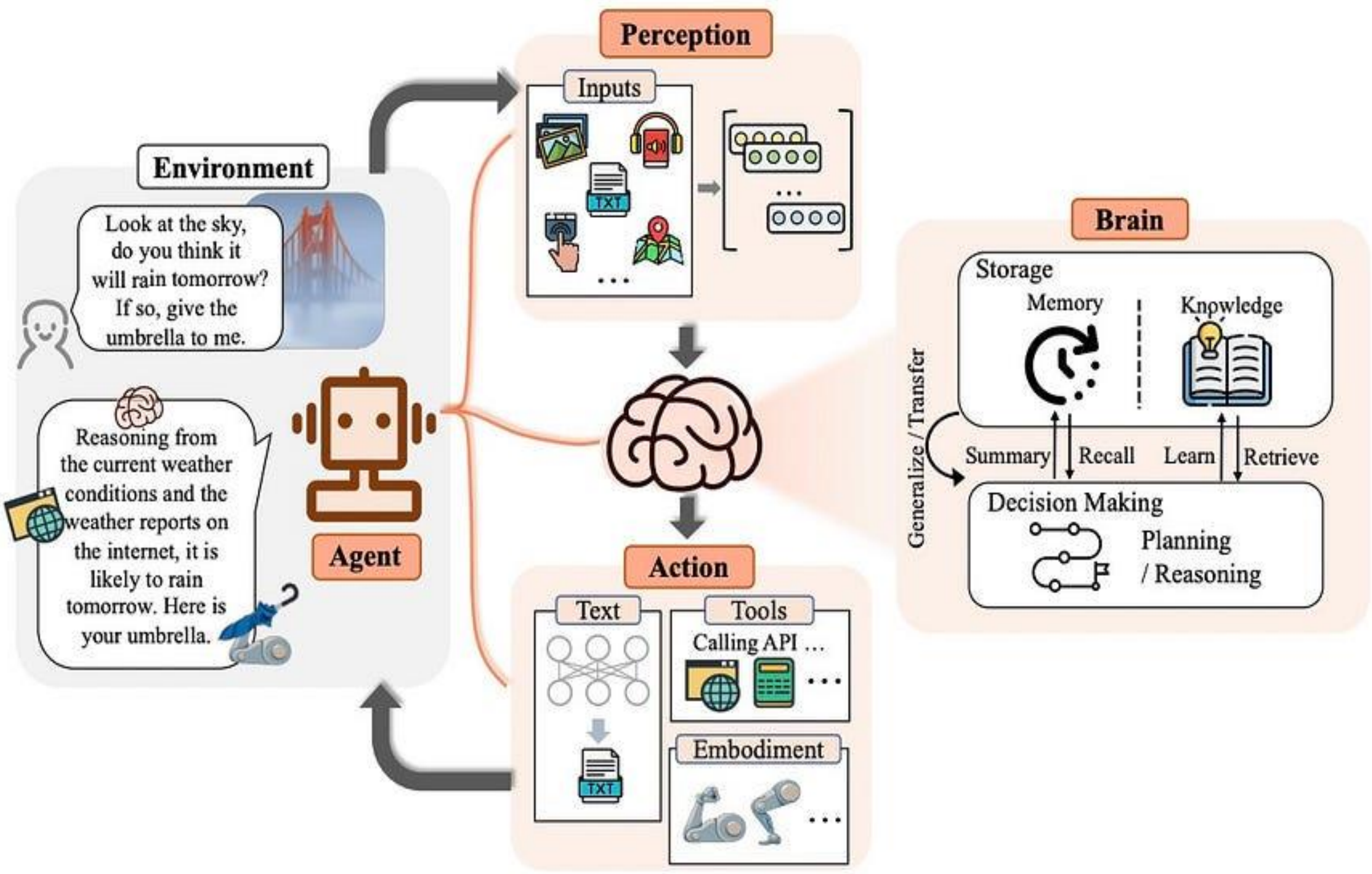
memory and thinking



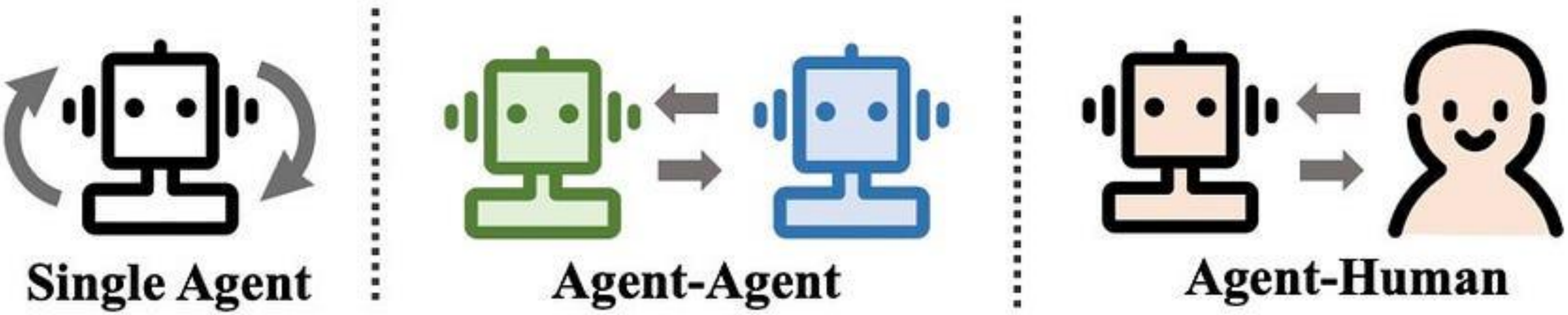
Artificial Intelligence

- A discipline of computer science used to create intelligent agents
- Intelligent Agents are systems that can learn, reason and act autonomously

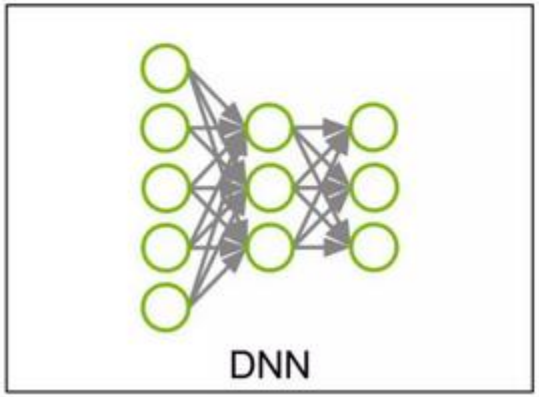
AI Agents


Environment
Look at the sky, do you think it will rain tomorrow? If so, give the umbrella to me.
Reasoning from the current weather conditions and the weather reports on the internet, it is likely to rain tomorrow. Here is your umbrella.
Agent
Perception
Inputs
TXT
...
Brain
Storage
Memory
Knowledge
Summary
Recall
Learn
Retrieve
Generalize / Transfer
Decision Making
Planning / Reasoning
Action
Text
TXT
Tools
Calling API ...
...
Embodiment
...

Agent Usage Scenarios

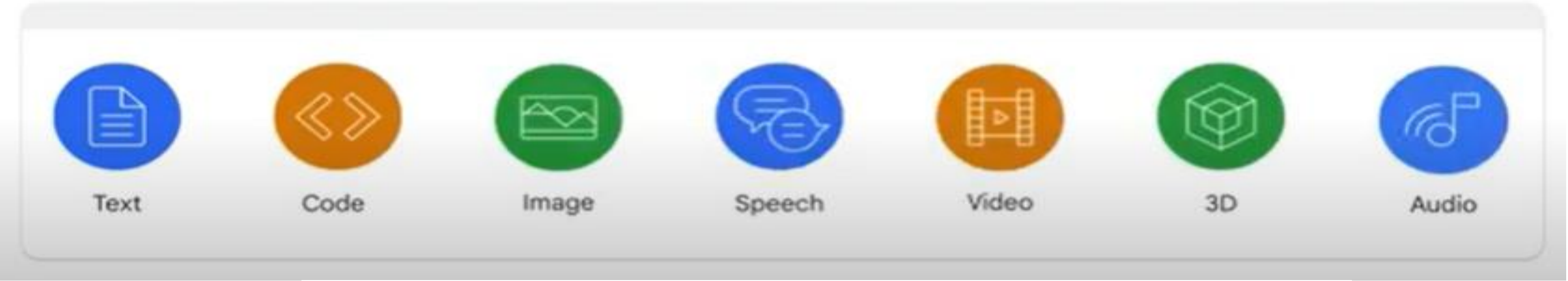


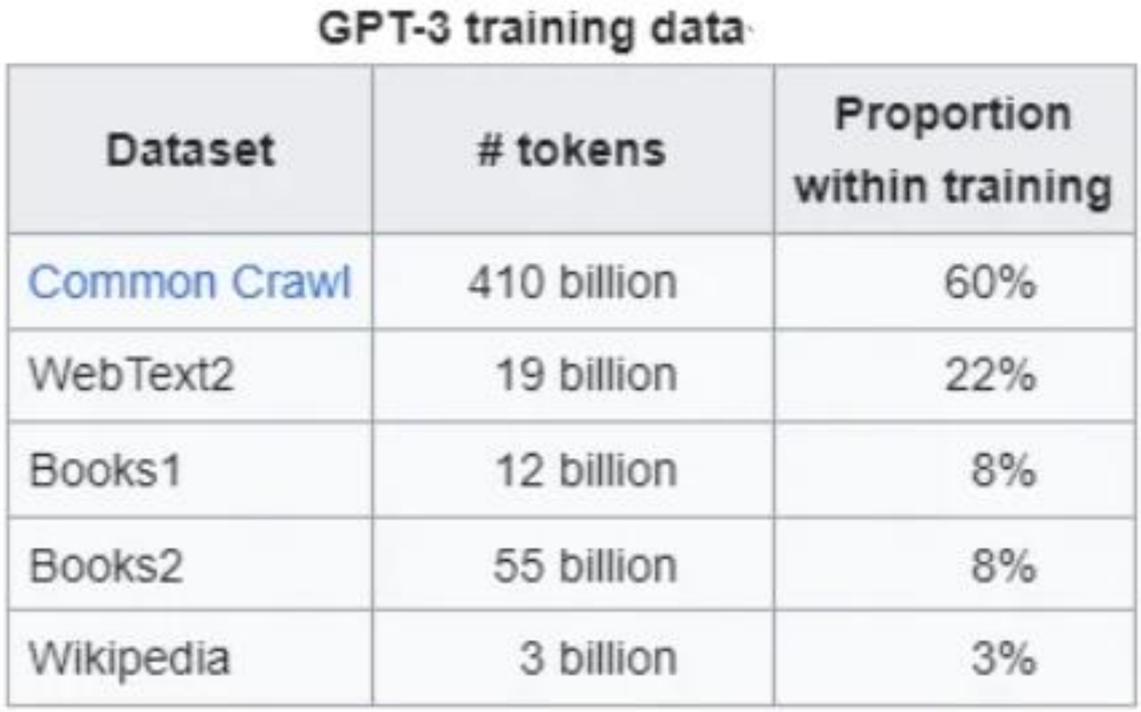
Fuel for AI Success



- BIG DATA
- Deep Models
- Computing Hardware Innovations

BIGDATA (Multimodal)





GPT-3 training data

Dataset	# tokens	Proportion within training
Common Crawl	410 billion	60%
WebText2	19 billion	22%
Books1	12 billion	8%
Books2	55 billion	8%
Wikipedia	3 billion	3%

GPU

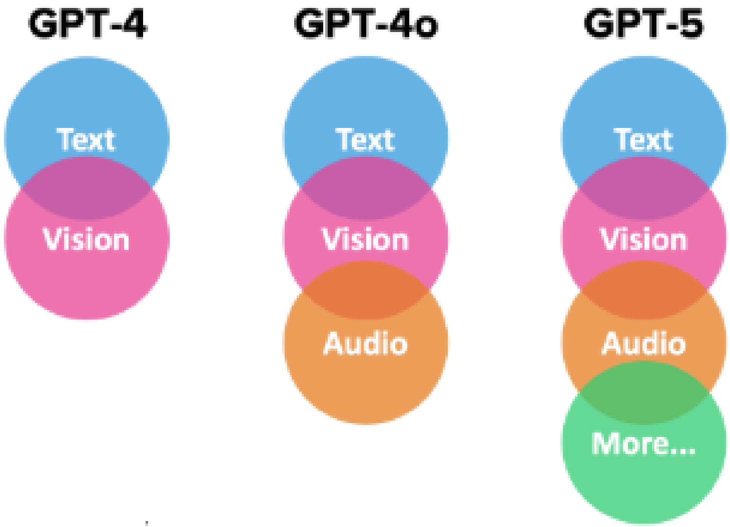
- NVIDIA H100 Architecture
  - 80B transistors
  - >15K cores
  - Optimized for parallel multiply/add operations
  - >1 TeraFLOPS at 16bit precision
  - >900 GB/s bandwidth

HOPPER H100 TENSOR CORE GPU

80B Transistors, TSMC 4N

Large Language Models (LLM)

Large Multimodal Models (LMM)



Gemini

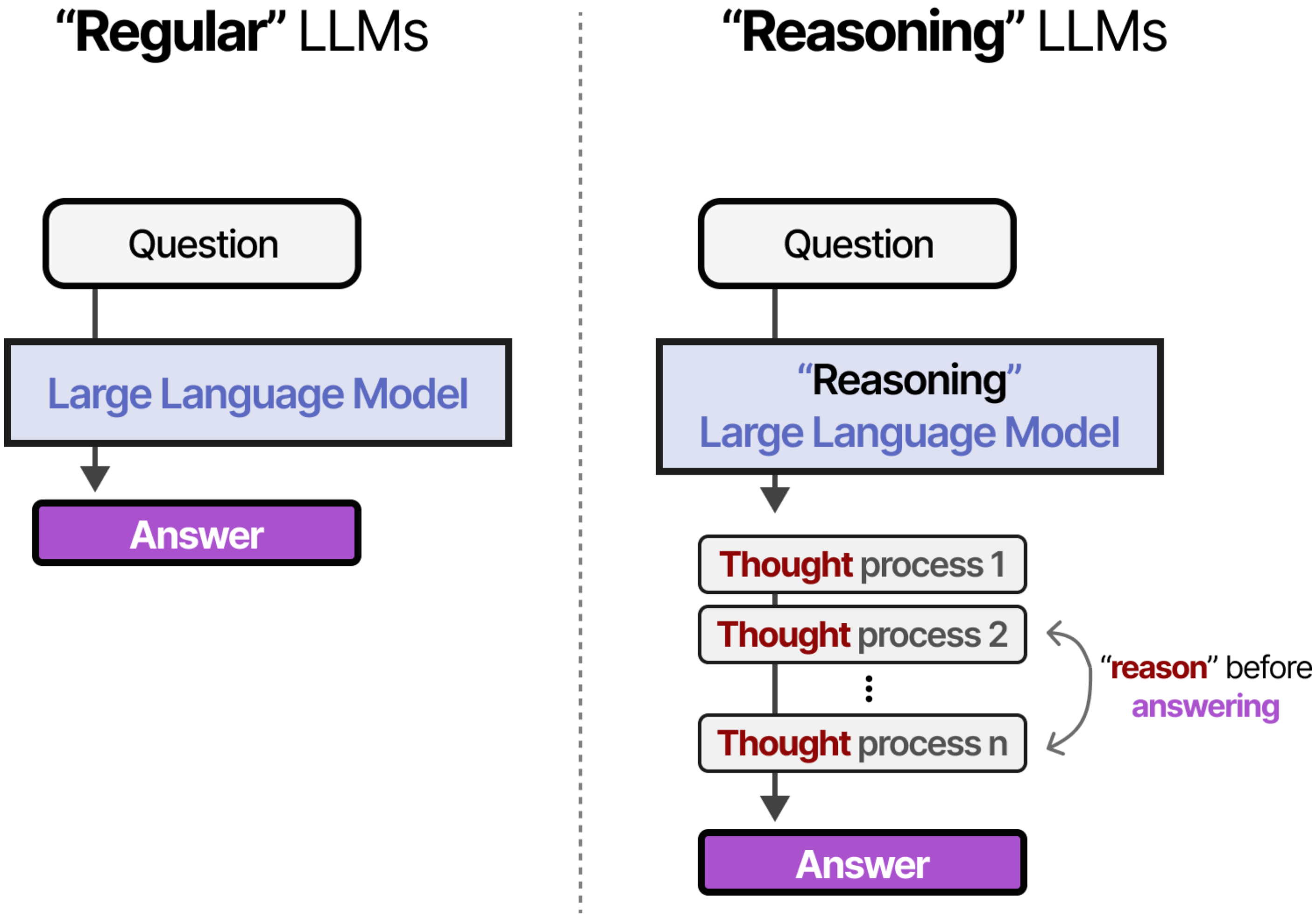
Large Reasoning Models (LRM)

DeepSeek-R1

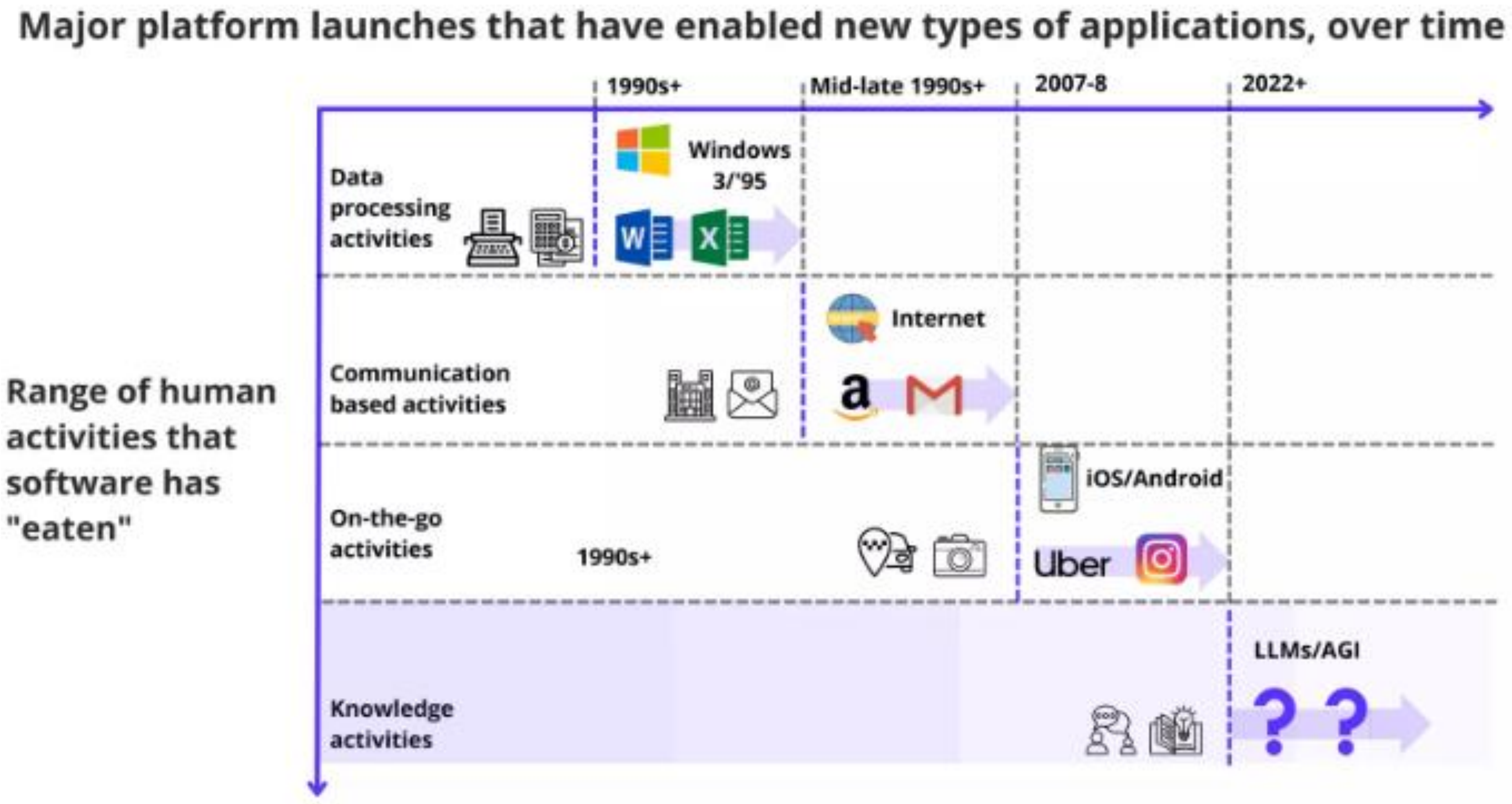
OpenAI o3-mini

Gemini 2.0 Flash Thinking

Large Reasoning Models (LRM)



App development breakthroughs


Major platform launches that have enabled new types of applications, over time
1990s+
Mid-late 1990s+
2007-8
2022+
Range of human activities that software has "eaten"
Data processing activities
Windows 3/'95
Communication based activities
Internet
On-the-go activities
1990s+
iOS/Android
Uber
Knowledge activities
LLMs/AGI
? ?

LLM/LMMs are the Engine &
AI Apps are the Product