

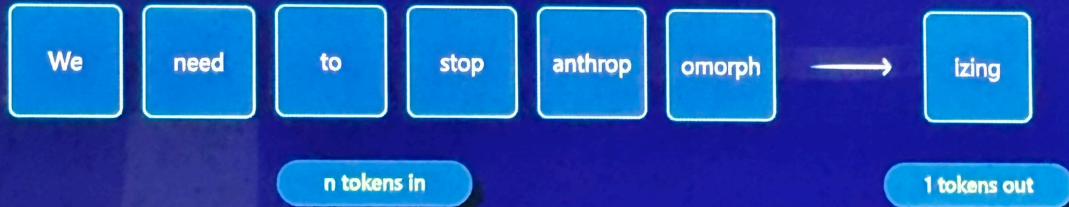
Week2

DataBiz팀 조용걸

ChatGPT

Language Model (LM)

랭귀지 모델 작동 방식



랭귀지 모델 작동 방식

in out

We need to stop

We need to stop anthrop

We need to stop anthropomorph

We need to stop anthropomorphizing

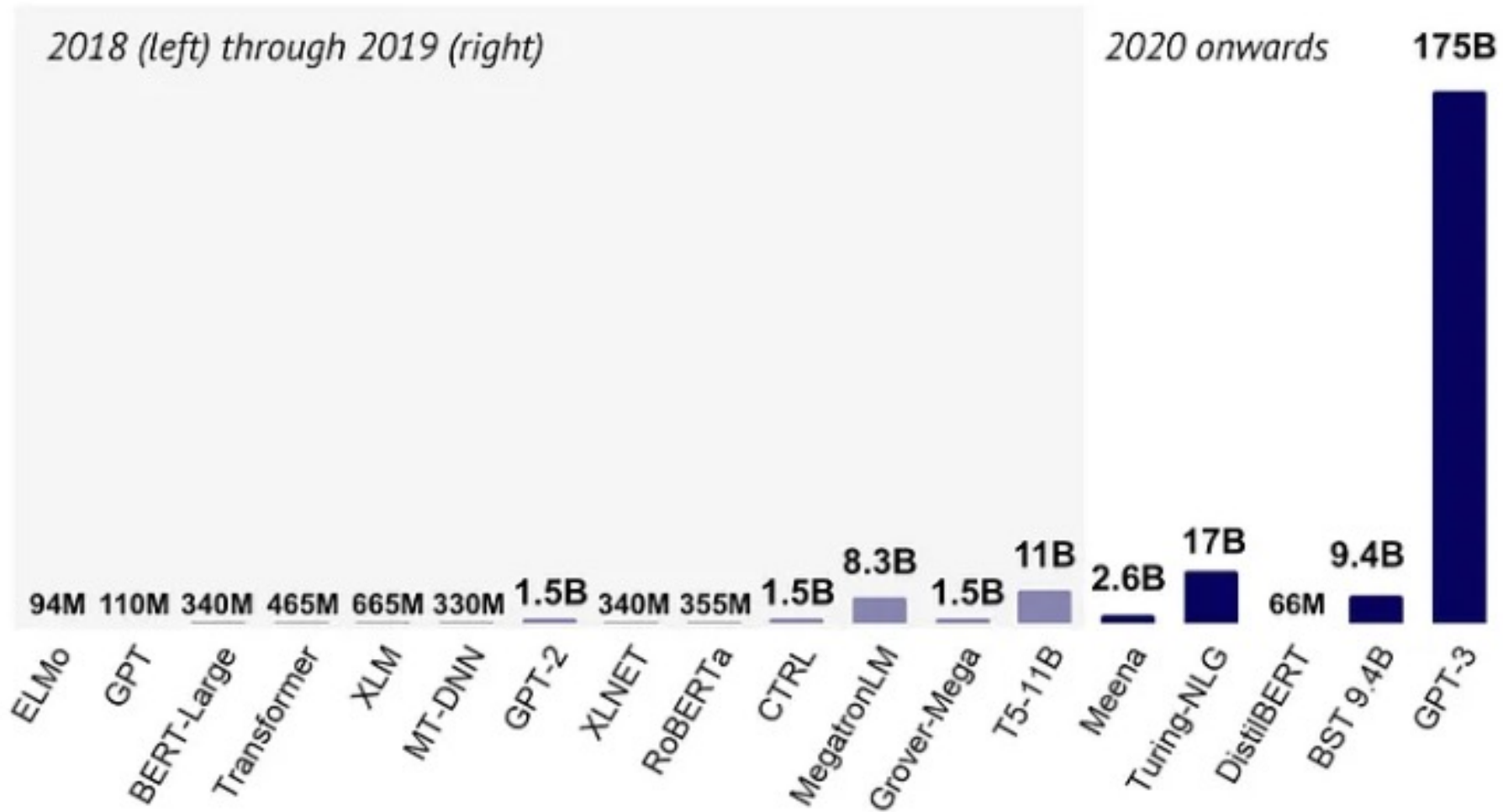
We need to stop anthropomorphizing Chat

We need to stop anthropomorphizing ChatG

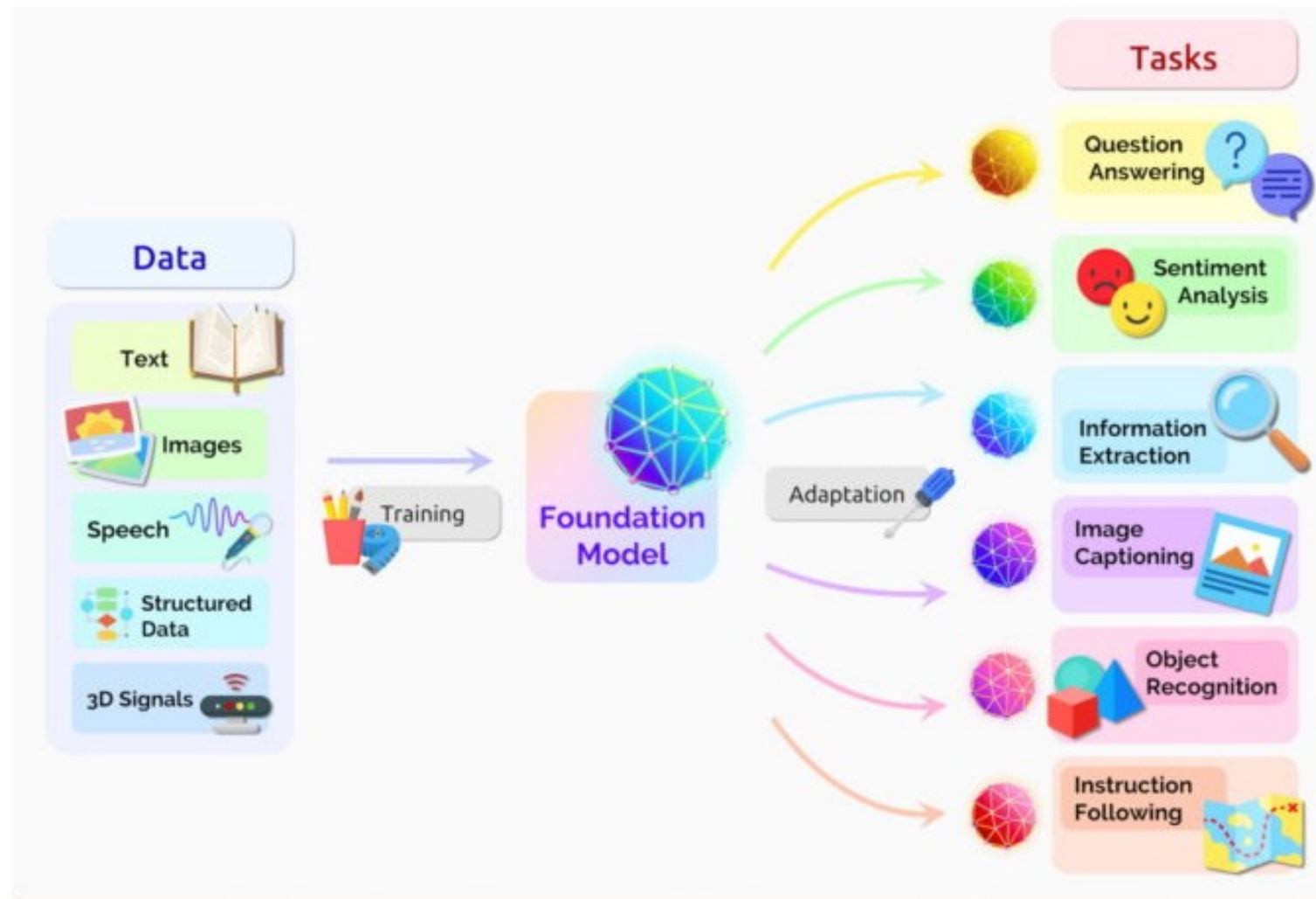
We need to stop anthropomorphizing ChatGPT

We need to stop anthropomorphizing ChatGPT.

Language Model history



Foundation Model의 등장



In a 2021 paper, researchers reported that foundation models are finding a wide array of uses.

Meta Llama 3 Instruct model performance

	Meta Llama 3 8B	Gemma 7B - It Measured	Mistral 7B Instruct Measured
MMLU 5-shot	68.4	53.3	58.4
GPQA 0-shot	34.2	21.4	26.3
HumanEval 0-shot	62.2	30.5	36.6
GSM-8K 8-shot, CoT	79.6	30.6	39.9
MATH 4-shot, CoT	30.0	12.2	11.0

	Meta Llama 3 70B	Gemini Pro 1.5 Published	Claude 3 Sonnet Published
MMLU 5-shot	82.0	81.9	79.0
GPQA 0-shot	39.5	41.5 CoT	38.5 CoT
HumanEval 0-shot	81.7	71.9	73.0
GSM-8K 8-shot, CoT	93.0	91.7 11-shot	92.3 0-shot
MATH 4-shot, CoT	50.4	58.5 Minerva prompt	40.5

<https://huggingface.co/yanolja/EEVE-Korean-Instruct-10.8B-v1.0>

데모

한계점

<https://platform.openai.com/docs/models/gpt-4-turbo-and-gpt-4>

1. 정보의 최신성

모델의 최신 Dataset 재학습

-> 결과 보장?

2. 할루시네이션 (hallucination)

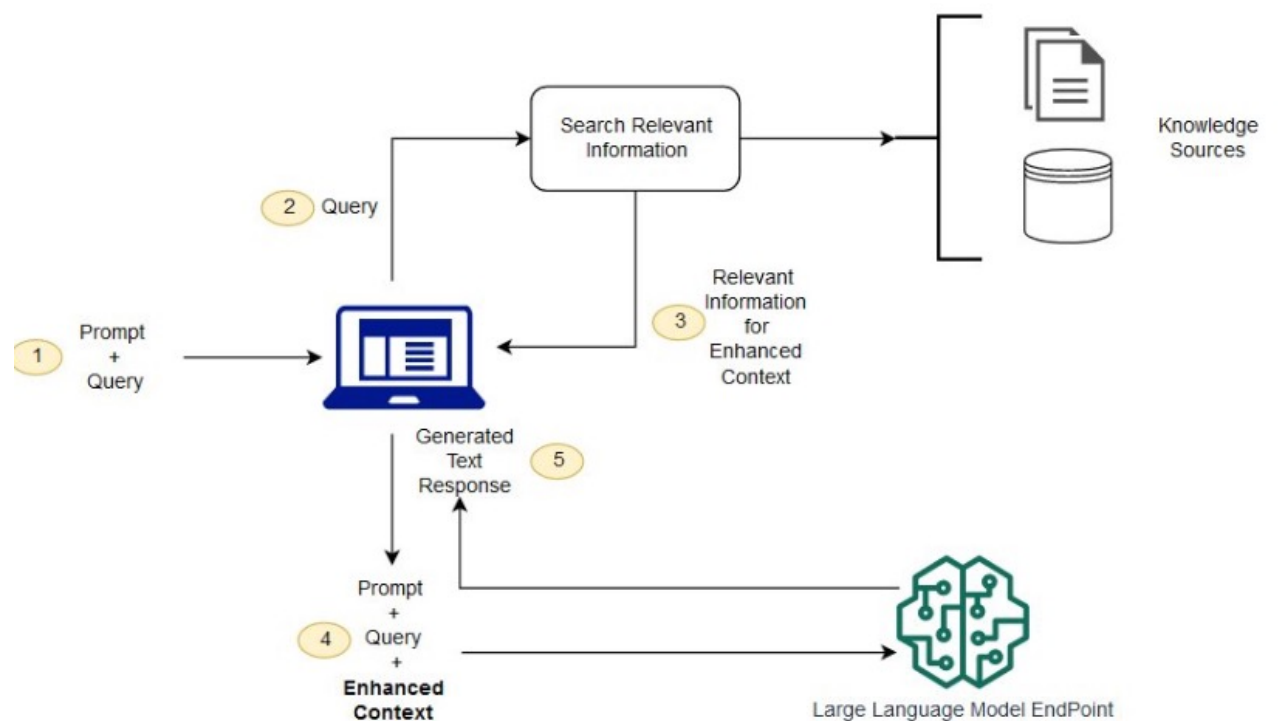
-> Resource(시간/장비 등)

> RAG

> Prompt Engineering

모델의 최신 Dataset 재학습
-> 결과 보장?
-> Resource(시간/장비 등)

Prompt Engineering
- RAG



서비스

데모

Prompt Engineering

Prompt engineering

<https://microsoft.github.io/Workshop-Interact-with-OpenAI-models/ko/>

감사합니다