

# **第2回**

# **大規模言語モデル**

# **Large language model**

---

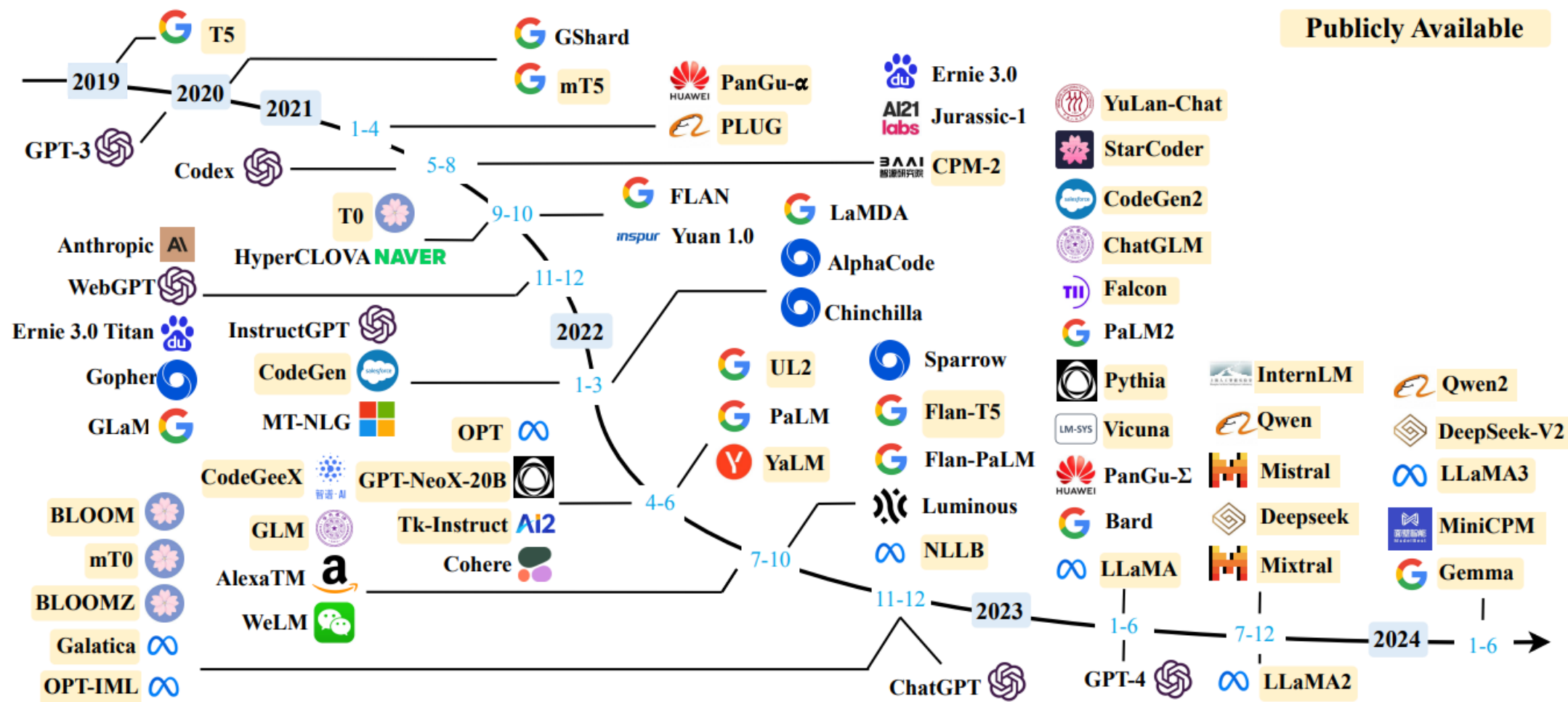
生成AI入門

# 自然言語処理（NLP）

---

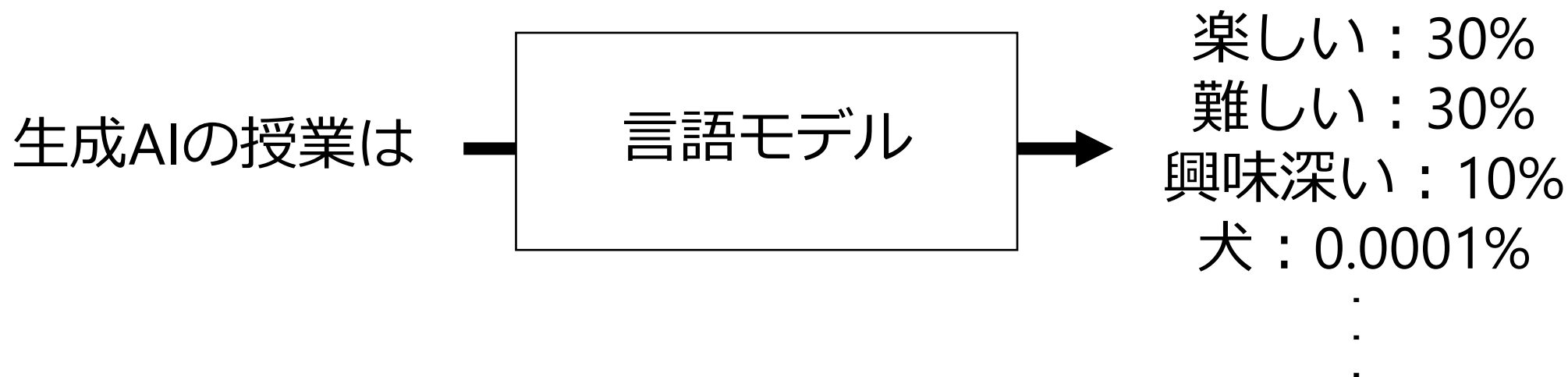
- コンピューターに自然言語を処理させること全般
  - 自然言語：日本語や英語など我々が普段使う言語
- 事例
  - 形態素解析：単語分解・品詞の判定する
  - 構文解析：修飾関係を解析するなどして文構造を把握する
  - 要約・パラフレーズ・意味理解
  - 機械翻訳
  - Chatbot

# 言語モデルマップ



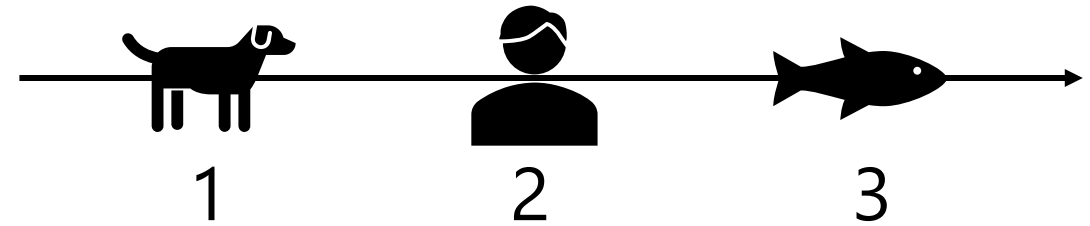
# (大規模) 言語モデル：(L)LM

- 単語列の生成過程を記述したもの
  - 言語における生成モデル
- 一般的に、近年のLMは単語列の生成確率をモデル化する



# 言葉を機械で扱うための表現を考える

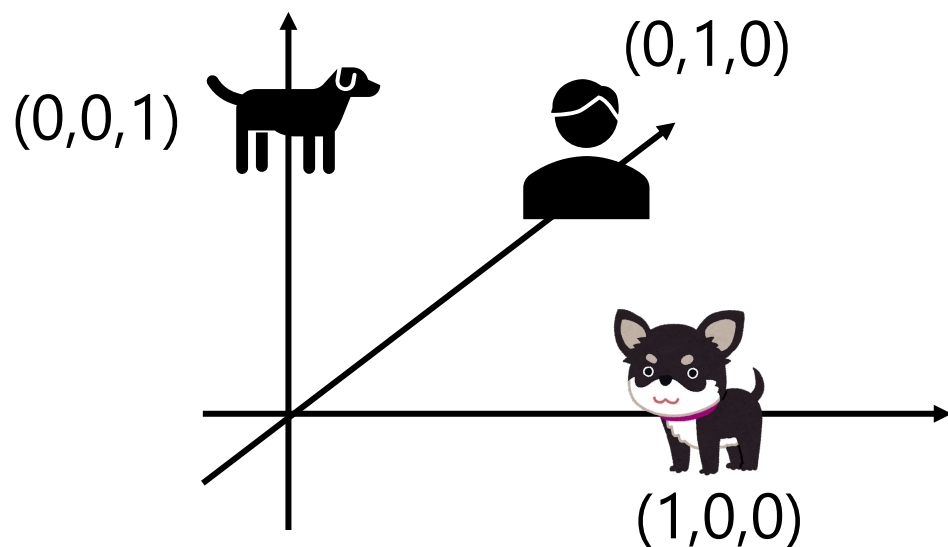
- “犬”は数値ではないので、機械で取り扱いにくい
- 犬 = 1, 男 = 2, 魚 = 3 のように数字を振る？
  - 量的データ vs. 質的データ
- チワワ = ?
  - 犬の仲間だが、どうする？



男 + 犬 = 魚？

# One-hot エンコーディング

- ある要素だけが1で、それ以外は0のベクトル表現
  - ベクトルの次元数は、単語数とする
  - 10000語で構成される言語では10000次元ベクトル

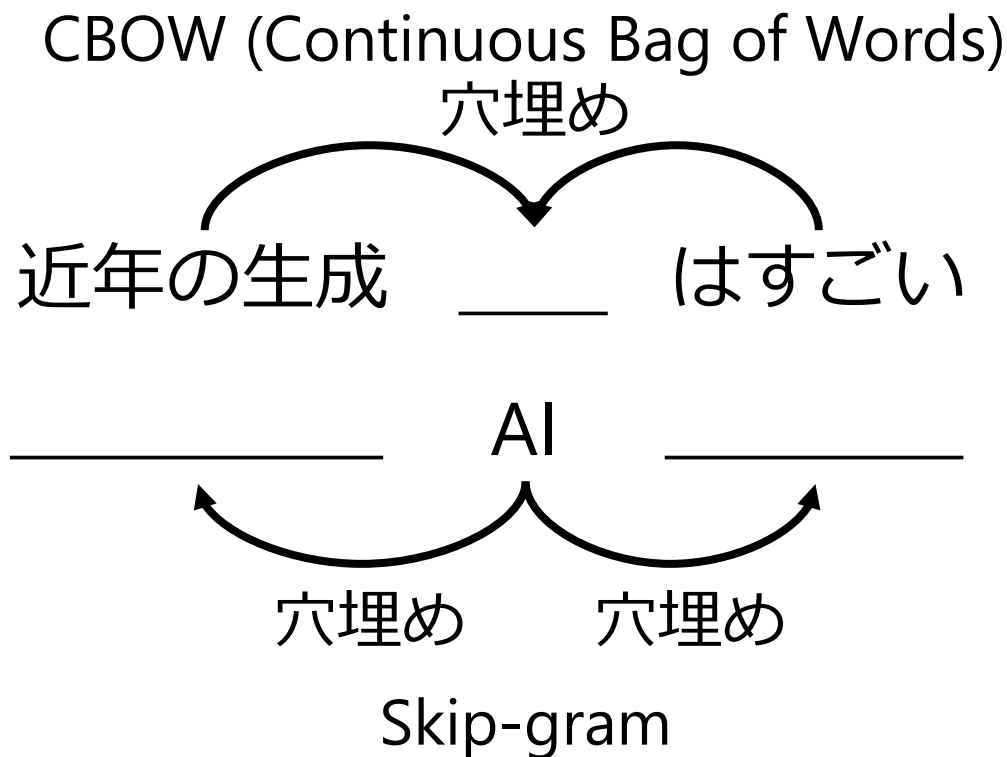


犬・男・チワワの3単語  
どれも直線距離は同じ

犬とチワワの意味の近さを  
考慮できない

# 単語間の関係性を考える

- ある単語からある単語を予測できたなら関係性があるはず



- 近年
- の
- 生成
- AI
- は
- すごい

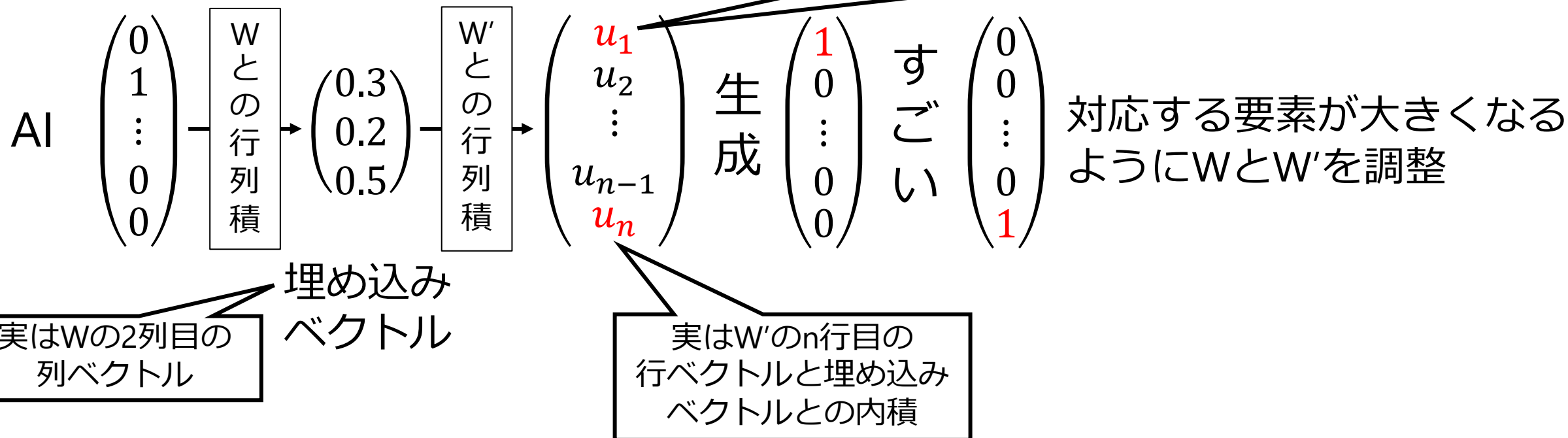
の単語間には関係がある

# Word2Vec

- 単語の関係性を持ったベクトルを学習する

- 埋め込み表現 (Embedding) ・ 分散表現

Skip-gram





# 埋め込みベクトル間の演算

- 共に現れやすい単語の埋め込みベクトルは類似
  - Word2Vecでは、共起する単語間を予測していた
- 意味を反映したような演算が可能
  - $\text{King} - \text{Man} + \text{Woman} = \text{Queen}$
  - 画像の埋め込みベクトルでも同様なことができる (DCGAN)
- 文章を考慮した埋め込みではない
  - Deep learning : 工学では深層学習の意味, 教育学では深い学習の意味
  - 文章中の意味によって、同じ単語でも埋め込みベクトルが異なるべき

# 文脈考慮の重要性

- 「それ」だけで予測は決まらない

それは\_\_\_\_\_

昨日は仕事がきつくて、それは\_\_\_\_\_

貴重なリンゴが売られている。それは\_\_\_\_\_

？

大変だったね

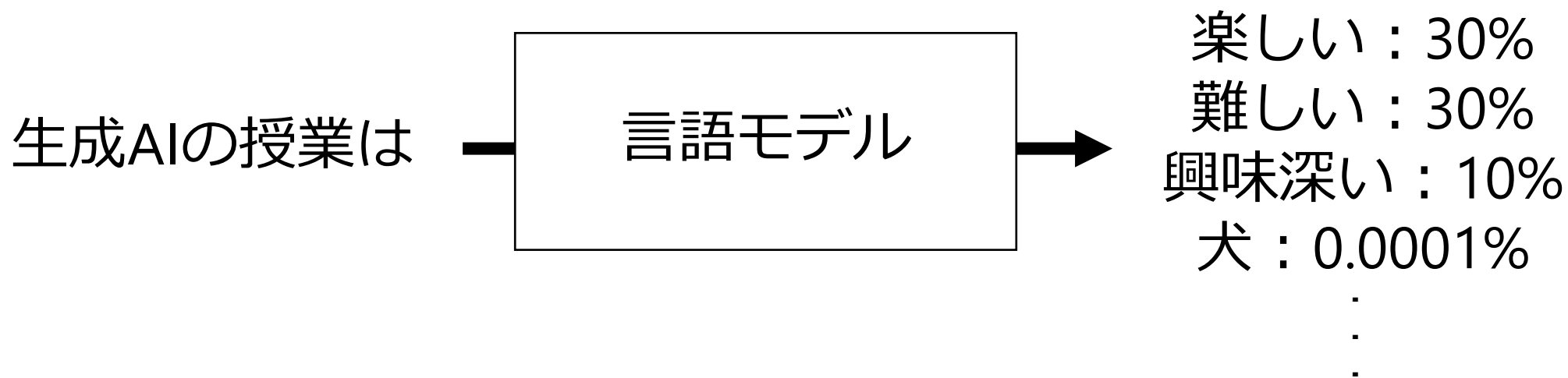
10万円する

赤色だ

何？

# 次単語予測

- 文章を入力に次の単語を予測
  - GPTの事前学習はこの学習方式を採用



# GPT-1を軽く読む

- Improving Language Understanding by Generative Pre-Training

## 3.1 Unsupervised pre-training

Given an unsupervised corpus of tokens  $\mathcal{U} = \{u_1, \dots, u_n\}$ , we use a standard language modeling objective to maximize the following likelihood:

今は単語と思ってよし

$$L_1(\mathcal{U}) = \sum_i \log P(u_i | u_{i-k}, \dots, u_{i-1}; \Theta)$$

調整パラメータ

(1)

自然な文章では  
この値は大きくなる

GPT

予測する単語

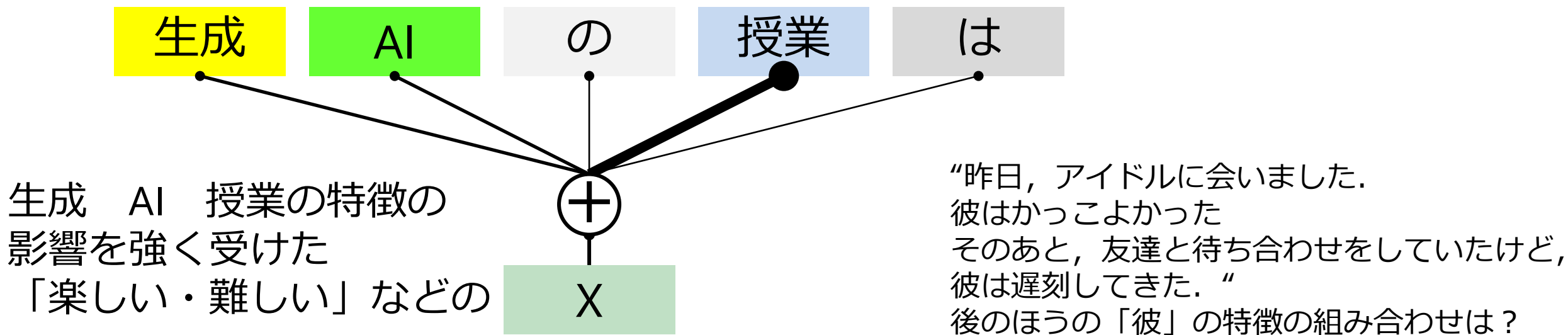
予測する単語より前  
に現れる単語列

自然な文章の例：「生成AIの授業は楽しい」「生成AIの授業は難しい」

- GPT(犬 | 生成)よりもGPT(AI | 生成)のほうがあり得る
- GPT(楽しい | 生成AIの授業は)とGPT(難しい | 生成AIの授業は)はどちらもあり得る

# GPTの実態

- 前に現れた単語の埋め込み特徴の組み合わせで次の単語の特徴を計算
  - TransformerのAttention機構
  - 予測する単語以外の特徴も同様の仕組みで計算される



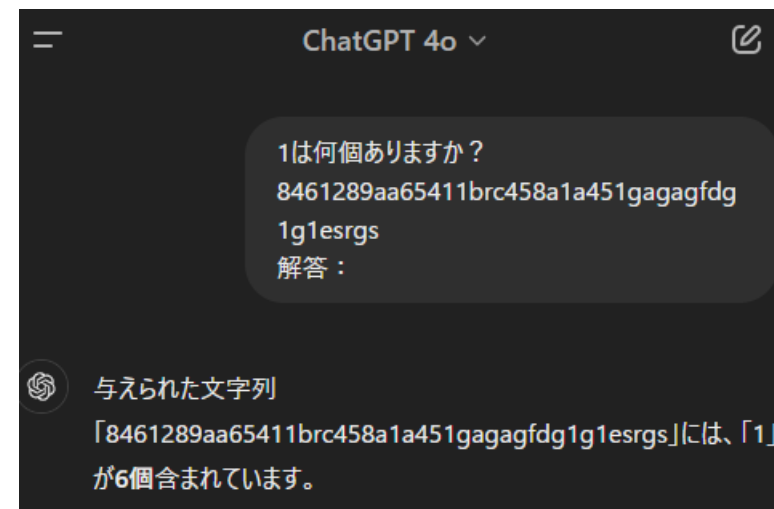
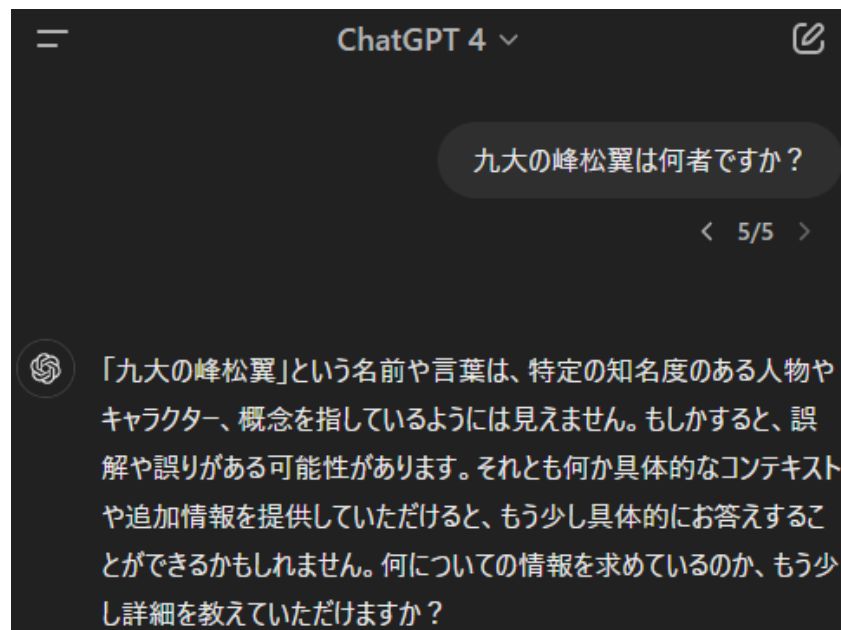
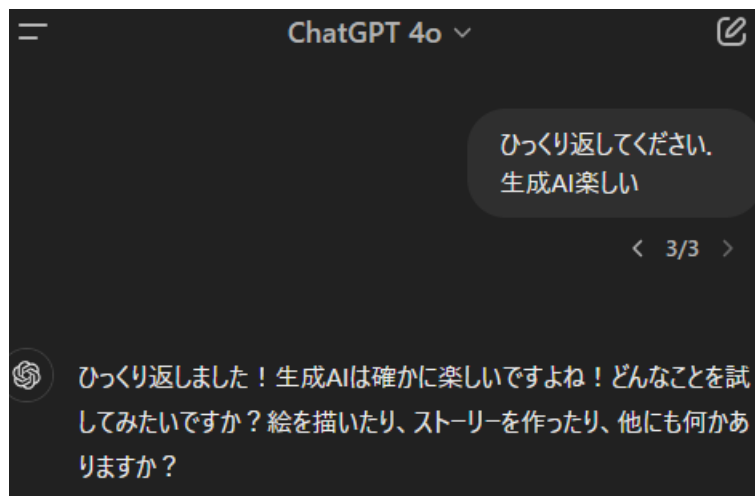
# GPTの歴史

- GPT-1 (2018)
  - 1.2億パラメータ, 7000冊の本
- GPT-2 (2019)
  - 15億パラメータ, 40GBの文書
  - 教師学習なしでも特定のタスクが解ける可能性を示唆
- GPT-3 (2020)
  - 1750億パラメータ, 570GBの文書
  - 例を示すと未学習タスクへ対応できる可能性を示唆
  - Few-shot, One-shot, In-context learning
- InstructGPT (2022)
  - 人間の好む文章に調整する手法
  - 強化学習・Alignment
- GPT-4 (2023), GPT-4o (2024)
  - 数千億～数兆パラメータ？ \*
- GPT-o1 (2024)
  - ??パラメータ
  - 論理的な展開を持つ文章を生成できるように訓練
  - Chain of Thought

# プロンプト

- LLMへの入力・指示

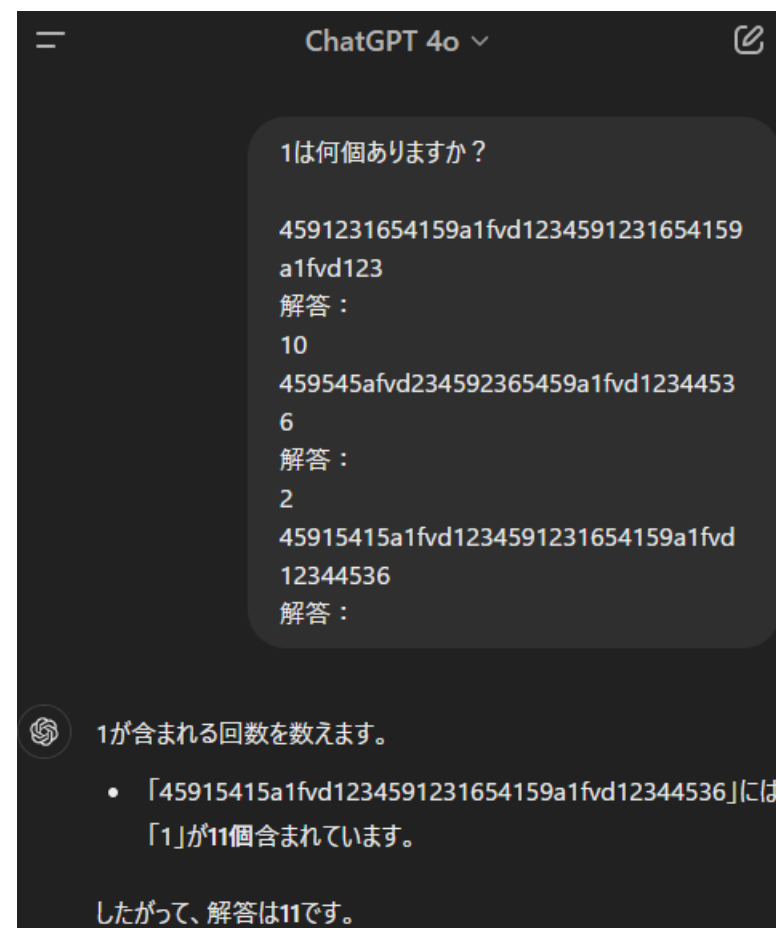
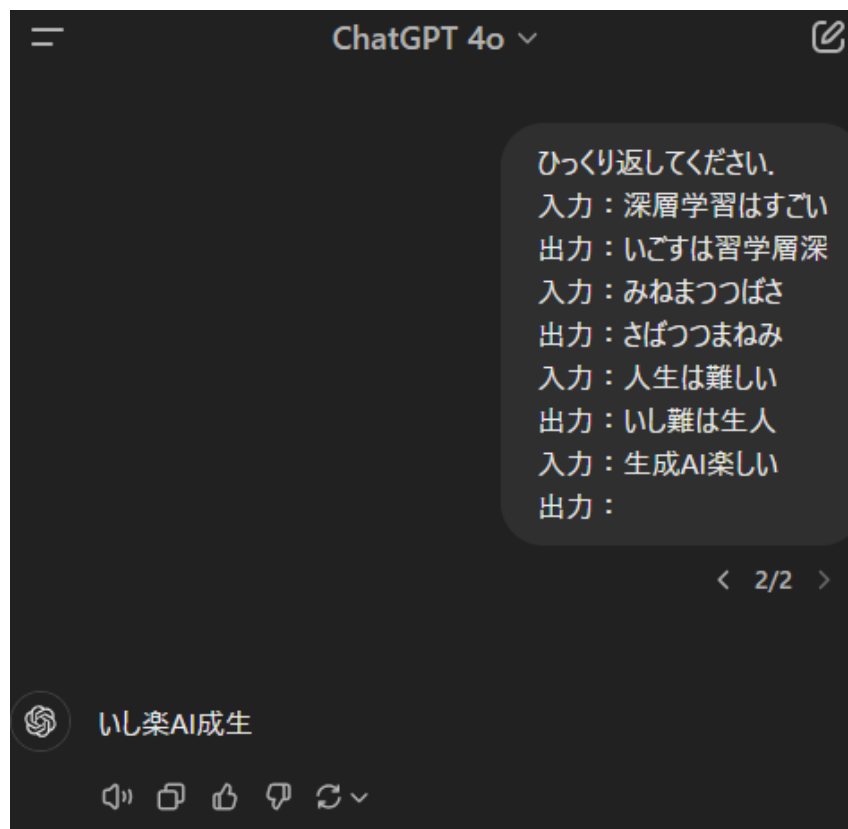
- LLMにうまく動いてもらうためにはよいプロンプトが必要なことが多い
- あいまいな指示では動いてもらえないのはなぜか？



# In-context learning

- 少数の事例を示して性能を改善する

- Few-shot, One-shot

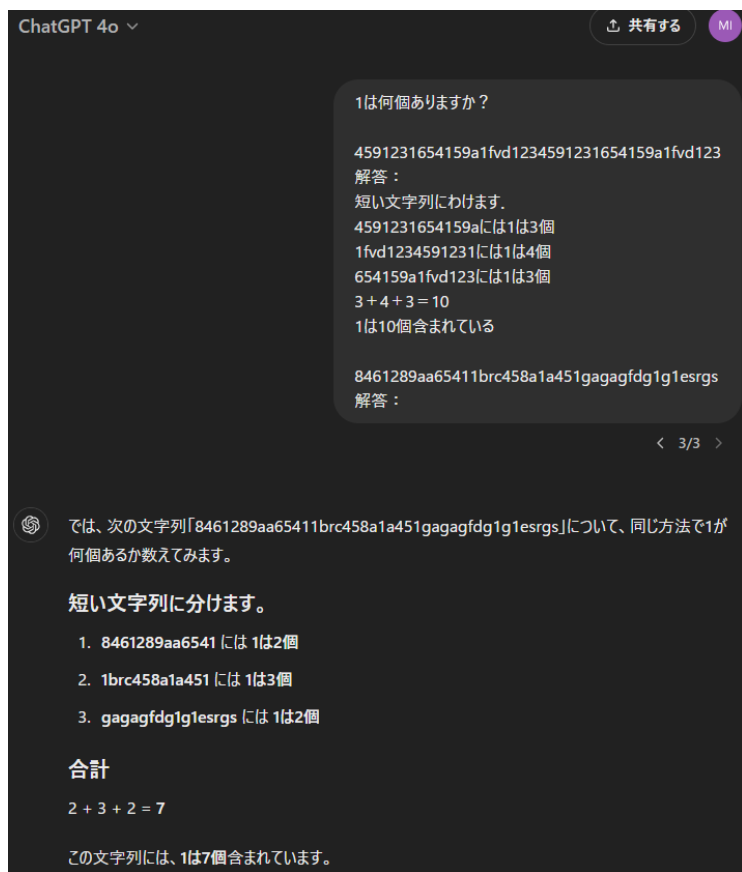




# Chain of Thought (CoT)

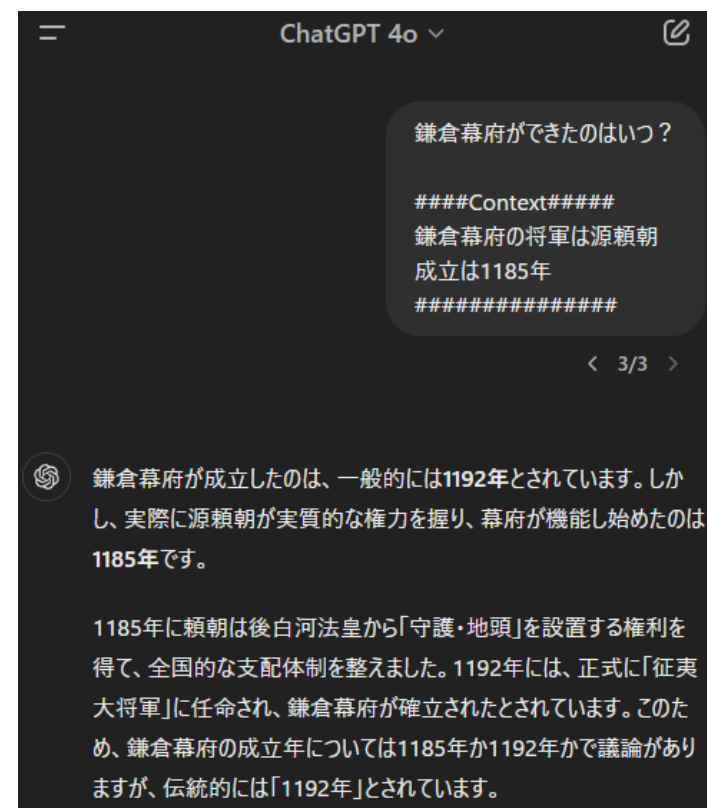
- 解答に至った経緯を事例として入力
- 解答に至った経緯を出力

- Zero-shot CoT



# Retrieval-Augmented Generation (RAG)

- 関係のある文書を検索等でプロンプトに組み込む
  - インターネット検索・独自のデータベース・知識グラフなどから情報を補足



# 関連話題

---

- Masked language model
  - BERT (Bidirectional Encoder Representations from Transformers)
- トークン
  - 文を一定の規則で区切ったのもで、単語単位になるとは限らない
- CLIP
  - 埋め込みベクトルは画像においても計算できる
- Prompt engineering
  - よいプロンプトを作る分野、promptの自動生成なども含まれる
  - CoTのほかにも、沢山の研究が行われている
- ReAct : SYNERGIZING REASONING + ACTING
  - LLMがタスクを解くためにすべきことを推論し、推論に基づいて行動し．．． を繰り返す

# 余談

