

生成AI

数理・データサイエンス・AI（リテラシーレベル）
モデルカリキュラム追加部分

社会で起きている変化

生成AI

①概要とその原理

生成AIとは

- 新しいデータやコンテンツ（画像、テキスト、音声など）を自動的に生成するAI技術



- 「どのようなデータを生成したいか」を**言葉で指示**できる生成AIも登場
 - 画像生成AIに対する指示の例：「食べかけのリンゴの画像を生成して」
 - テキスト生成AIに対する指示の例：「AIの原理の説明文を生成して」
 - これら指示文は**プロンプト(prompt)**と呼ばれる

生成AIの例 (1/2)

画像生成の例

- “Make our society happy with data science” (我々の社会をデータサイエンスで幸せに!) というプロンプトに対する画像生成結果



StableDiffusionWeb.com



ChatGPT (OpenAI社, ChatGPT 4o)

生成AIの例 (2/2)

テキスト生成の例

- 「AIの原理を4文ぐらいで説明してください」というプロンプト（質問）に対するテキスト生成結果

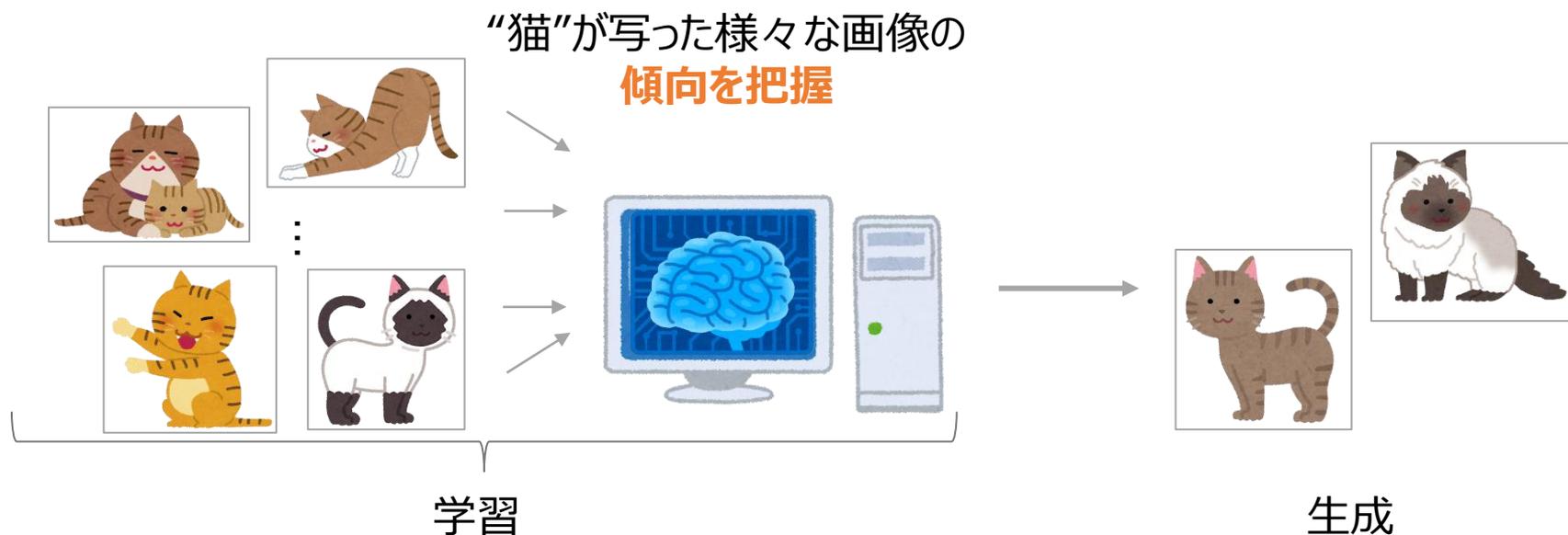


AI（人工知能）は、大量のデータを学習してパターンや規則を見つけ出し、その知識をもとに推論や判断を行います。これには、ニューラルネットワークや機械学習アルゴリズムが使われ、特に深層学習が近年注目されています。AIは訓練データを用いてモデルを構築し、新しいデータに対しても予測や分類を行うことができます。また、反復的な学習プロセスを通じて、自己改善を行い精度を高めていきます。

ChatGPT (OpenAI社, ChatGPT 4o, 2024年8月生成)

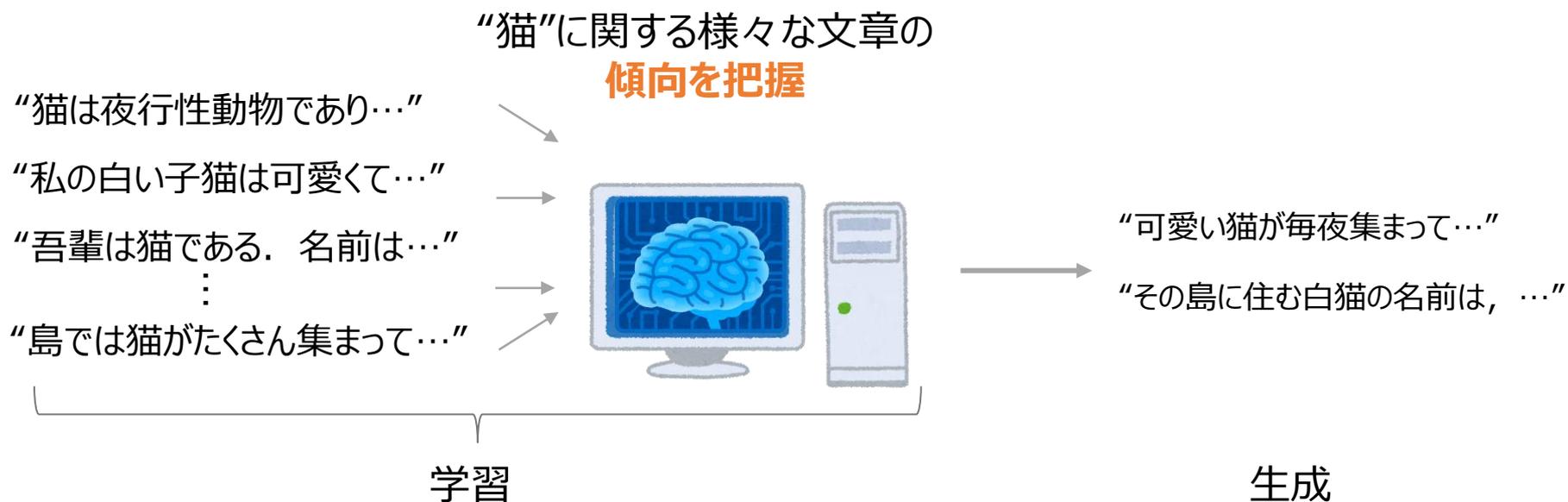
生成AIの原理 (1/3)

画像生成AIの場合



生成AIの原理 (2/3)

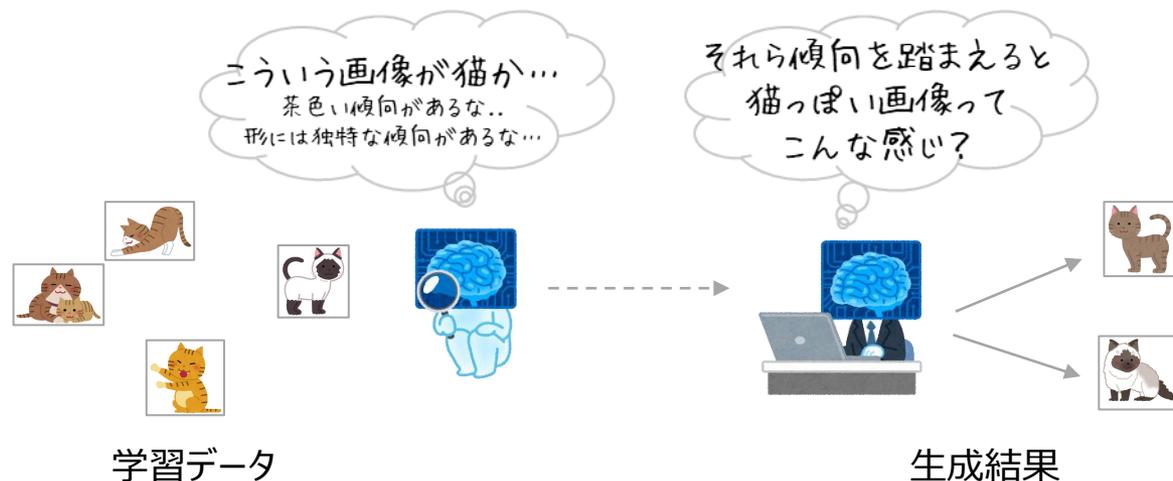
テキスト生成AIの場合



生成AIの原理 (3/3)

まとめると...

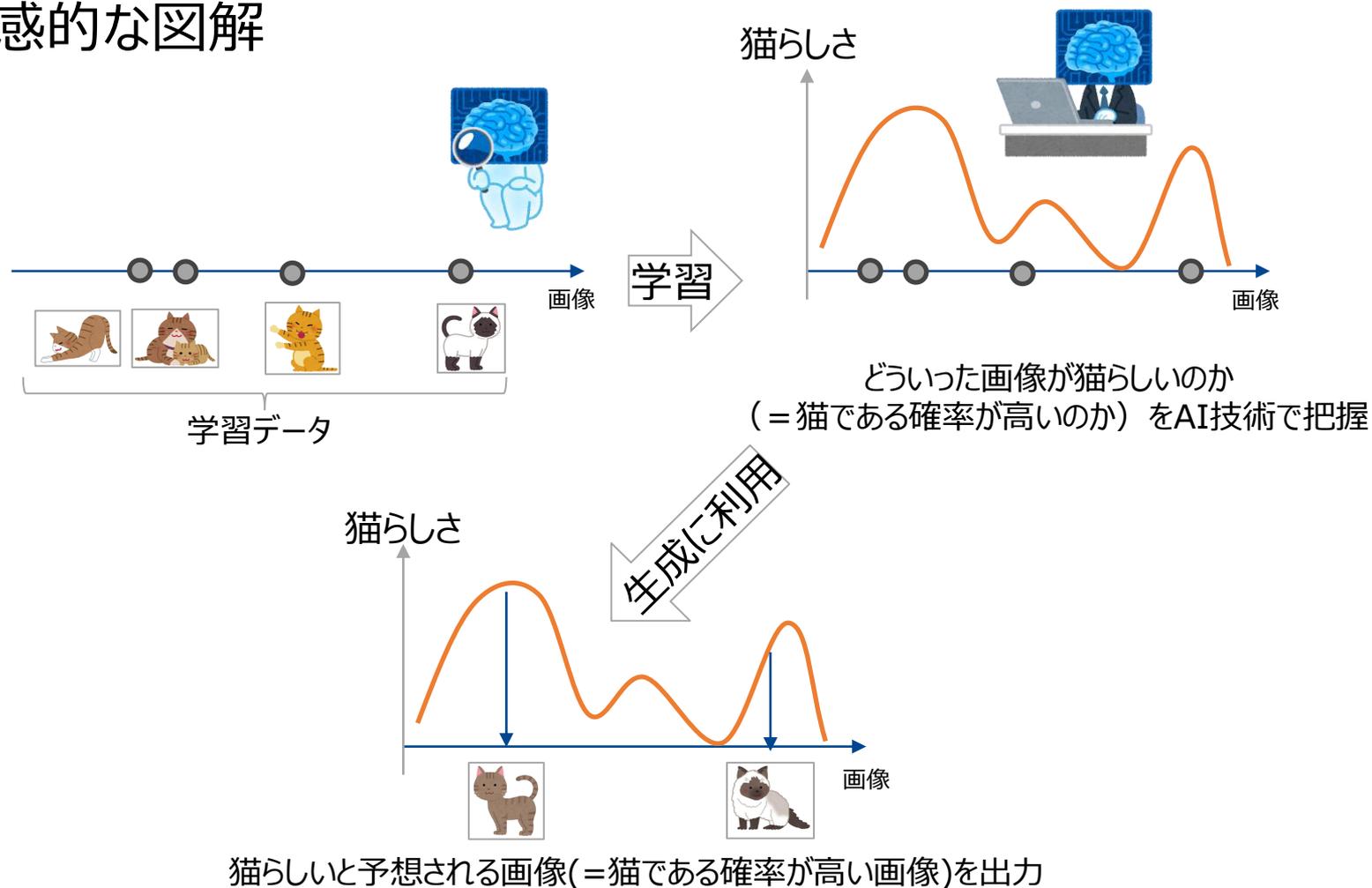
- 多くのデータを学習し、「それらとは異なるが、似たデータ」を作り出す
 - 例えば、1万人分の顔画像で学習した生成AIは、「その1万人とは異なるが、実際にいそうな人の」顔画像を、無限に作ることができる
- 従って、「学習したデータの背後に潜む傾向」を把握するAI



参考：生成AIの原理

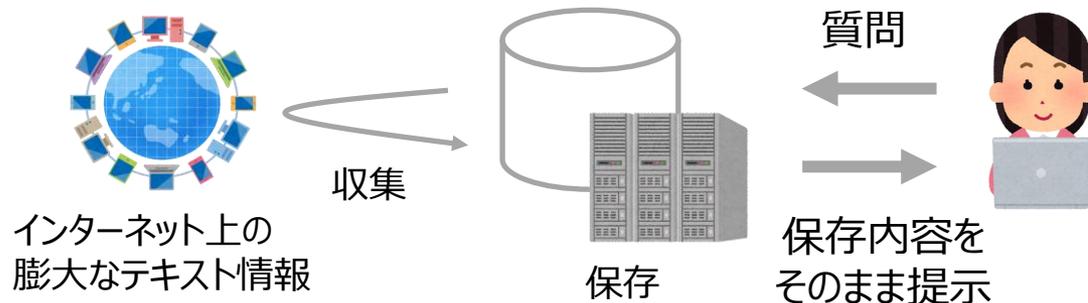
正確には「学習したデータの確率分布を推定するAI」

● 直感的な図解

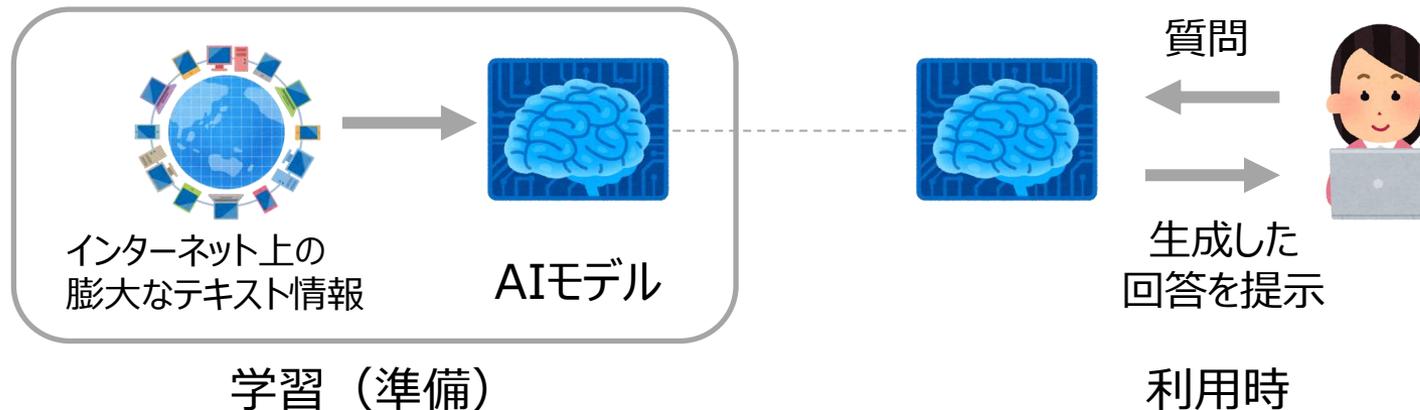


「検索」と「生成AI」の違い： テキストの場合

● 検索



● 生成AI



● 最近「検索」も併用する生成AIモデルも登場

- 検索拡張生成(Retrieval Augmented Generation, RAG)と呼ばれる

社会で起きている変化

生成AI

②最近の技術的發展

生成AIの発展

- AIの理論・技術の進化に伴い，生成AIも劇的に発展
 - 深層ニューラルネットワークが使われるようになった2010年代半ばから
 - 特に2020年代になってからの発展が著しい
- 3方向の発展
 - **生成対象**の発展
 - より様々なものが生成可能に
 - **生成方式**の発展
 - 「とにかく生成」から「特定の結果を目指した生成」まで
 - **生成結果のクオリティ**の発展
 - AIによる生成物なのか本物なのか，人間にも区別がつかないレベルに

生成対象の発展

- テキスト
 - 詩歌や小説, 質問に対する回答, プログラムコードなど様々
- 静止画
 - 現実を写したようなものからイラストや線画のようなものまで様々
- 動画
 - 次スライド
- 音 (音声・音楽・環境音)
- ロボットの動作生成



動画生成AIの例



プロンプト: A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage. She wears a black leather jacket, a long red dress, and black boots, and carries a black purse. She wears sunglasses and red lipstick. She walks confidently and casually. The street is damp and reflective, creating a mirror effect of the colorful lights. Many pedestrians walk about.

<https://openai.com/index/sora/>

Gifファイル : <https://www.gizmodo.jp/2024/02/openai-video-generator-sora-is-breathtaking-terrifying.html>

生成方式の発展： 生成物についてより細かい制御が可能に

- 学習したデータに似たデータをランダムに生成
 - 顔画像生成ならば：とにかく人間の顔をランダムに生成
- 特定のデータに似たデータを生成
 - 顔画像生成ならば：A氏のある顔画像を与えると、別の表情・角度でのA氏の顔画像を生成
- プロンプト（指示文）に従ってデータを生成
 - 顔画像生成ならば：「眼鏡をかけた25歳の日本人女性の笑った横顔」というプロンプトどおりに顔画像を生成
- 後二者のハイブリッド（=与えたデータをプロンプトの指示に従って加工する方式）も存在

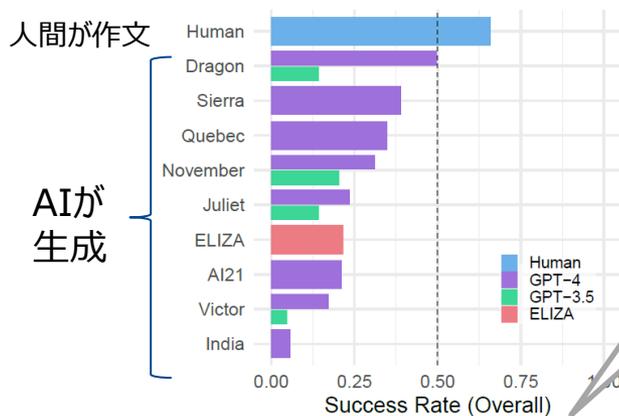
生成結果のクオリティの発展 (1/2)

テキスト生成

- 「人間が作文したもの」と「AI生成文」の違いを見つけるのは、もはや困難なレベル
 - 2023年に行われた対話実験では、「生成AI (GPT-4) による文」の約半分は「人間による作文」と (人間に) 判断された

Does GPT-4 pass the Turing test?

Cameron R. Jones and Benjamin K. Bergen
 UC San Diego,
 9500 Gilman Dr, San Diego, CA
 {cameron, bkbergen}@ucsd.edu



人間に
「人間による文だ」
と思わせた割合

- ちなみに「人間による作文」の66%は「人間による作文」と (人間に) 判断された
 - なので、現時点では、わずかだが「人間による作文」のほうが人間らしさを保っている
- 「人間 or AI」を識別するAI (AI-Content detector)も存在
 - しかし、上記の困難性により、精度は不十分

生成結果のクオリティの発展 (2/2)

画像生成

- この10年で格段に進化

特定物の生成
例：顔



2014



2015



2016



2017



2018

<https://www.iangoodfellow.com/slides/2019-05-07.pdf>

プロンプトによる任意画像の生成
例：「アボカドの形をした椅子」



<https://www.technologyreview.jp/s/230408/this-avocado-armchair-could-be-the-future-of-ai/>

動画の生成



<https://openai.com/index/sora/>

- 現時点(2024)では細かなところで不自然な部分が残るケースも



<https://openai.com/index/sora/>



文字らしくはあるものの、読めない

- ただし進化は続いており、今後真実との見分けはますます困難に

生成AIの活用事例

● テキスト生成AI

- 質問回答（対話）
- テキストコンテンツ生成
- 翻訳・要約・推敲（よりよい文章にする）



● 画像生成AI

- 静止画・動画コンテンツ生成
- グラフィックデザイン支援
- 画像補正・加工



商用ソフトウェア・サービスとして、ビジネス現場等で実利用されているものもある

● 音生成AI

- テキスト読み上げ・アナウンス生成
- 音楽加工
- 歌詞からの楽曲生成



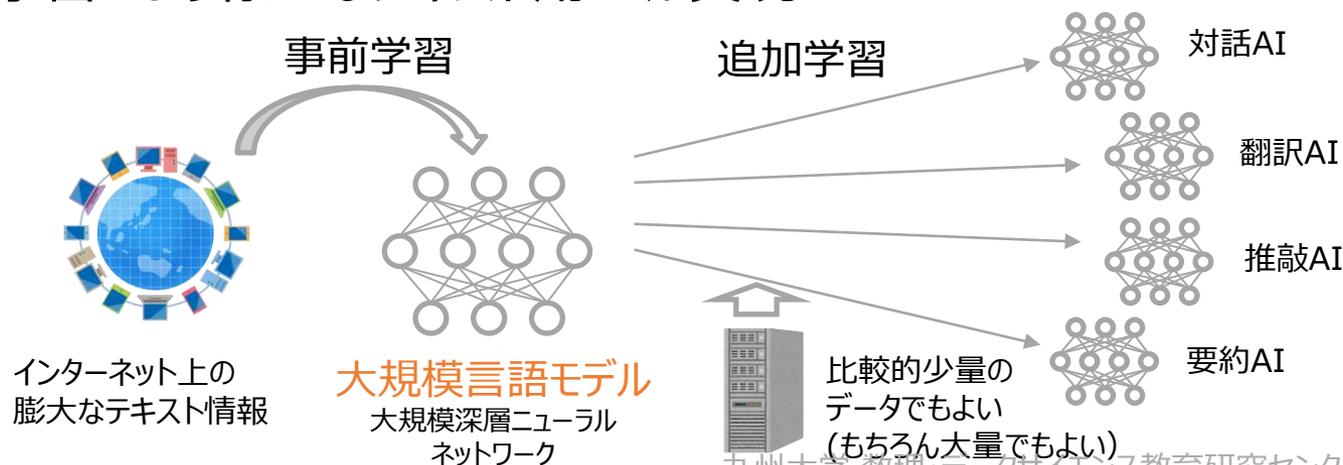
参考：基盤モデルと大規模言語モデル（1/3）

● 基盤モデル

- 「膨大な」データで事前学習された「大規模な」深層ニューラルネットワーク
- 各自が実現したい課題のデータを使って基盤モデルを追加学習することで、比較的容易に（例えば、その課題用のデータはそこまで多くなくても）高精度なAIを実現可能

● 大規模言語モデル（Large Language Model, LLM）

- テキストデータを対象とした基盤モデル
- 追加学習により様々なテキスト用AIが実現



参考：基盤モデルと大規模言語モデル (2/3)

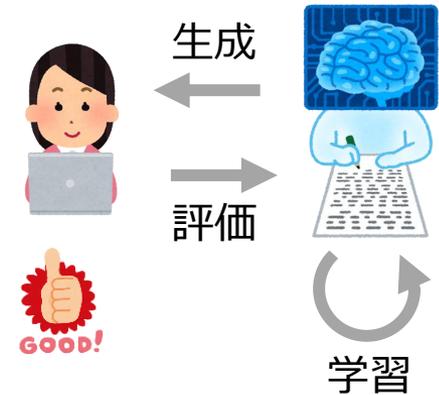
どれぐらい大規模なのか？

- 大規模言語モデルGPT-3 (OpenAI社, 2020年)
 - 1750億のパラメータ (学習によって設定されるべき変数の数) を持った深層ニューラルネットワーク
 - 学習には570GB (ギガバイト) のテキストデータを利用
 - 2022年の英語版Wikipediaの全データが約100GBぐらい
 - それよりも5倍以上多いデータを使って事前学習！
- 大規模言語モデルGPT-4 (OpenAI社, 2022年)
 - パラメータ数やデータ量の詳細は公開されていない
 - 1台あたり約100万円のGPUを25,000台用いて90日かけて学習？？
 - <https://www.ikangai.com/the-secrets-of-gpt-4-leaked/>

参考：基盤モデルと大規模言語モデル (3/3)

人間の力を借りて高精度化

- LLMを追加学習して「対話生成AI」を作る場合
 - 当然「適切な対話」を生成できるようにしたい
 - しかし「どんな対話が適切か？」を数式で表現するのは困難
 - ならば「人間に評価してもらいつつ、より高評価になるように学習しよう！」



- 「人間のフィードバックによる強化学習」と呼ばれる
 - Reinforcement learning from human feedback (RLHF)
 - GPTを追加学習して対話生成AI「ChatGPT」を作る際にも利用

データ・AIの活用領域

生成AIの応用

様々な生成結果例

テキスト生成AIの応用例 (1/6)

対話 (チャット)

- 相手の入力テキストに対する返答テキストを生成するAI
- 顧客等の質問に自動回答するサービスにも利用可能
 - 右の例は、市役所の窓口役として質問に回答した例

回答

質問

住民票を新住所に移す時には、どのような手続きが必要でしょうか？



住民票を新住所に移す際には、以下の手続きが必要です。これは「転出届」と「転入届」、「転居届」の手続きに分かれます。

1. 転出届

- 対象: 現在住んでいる市区町村外に引っ越しする場合。
- 手続き場所: 現在住んでいる市区町村の役所。
- 必要なもの:
 - 本人確認書類 (運転免許証、パスポートなど)
 - 印鑑 (場合によっては不要なこともあります)
- 手続き内容:
 - 引っ越しの予定日までに転出届を提出し、転出証明書を受け取ります。この証明書は新しい住所での転入手続きに必要です。

2. 転入届

- 対象: 新しい市区町村に引っ越してきた場合。 ...

ChatGPT (OpenAI社, ChatGPT 4o, 2024年8月生成)

- 各企業の業務に特化した対話へのカスタマイズも可能
 - 各企業が集積したノウハウを学習
 - 「知恵袋」として利用でき、その企業の業務の効率化に貢献

テキスト生成AIの応用例 (2/6)

テキストコンテンツ生成

「データサイエンス」という言葉を使った面白い話を短く作ってください

 データサイエンティストが友人にこう言ったんです。

「僕の恋愛も、データサイエンスで解析すればうまくいかもと思ったんだ。」

友人が聞き返します。

「で、どうだったの？」

データサイエンティストはため息をついて答えました。

「残念ながら、相関はゼロだったよ。」

友人が笑いながら一言。

「それ、データが足りなかったんじゃない？」

ヒットしそうな曲の歌詞を「データサイエンス」というフレーズをいれて作って

 もちろんです！ データサイエンスをテーマにしたポップソング風の歌詞を考えてみました。

タイトル: 「データのダンス」

(Verse 1)

データの海で 泳いでる
僕らの未来が キラキラと
無限の可能性 広がる空
分析のリズムで 踊り出そう

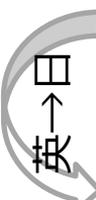
(Chorus)

データサイエンス 君と描く夢
数式の旋律 響かせて
ビッグデータの波に 乗りこなして
この世界を変えてく 未来のダンス ...

テキスト生成AIの応用例 (3/6)

翻訳

- 
 「データサイエンスは意外に面白い。人間が色々やっていることと関連しているのだ。例えば、データサイエンスの分野において『予測』は重要なトピックであり、高度な数学を用いることも多い。一方、我々人間も予測を行っている。子供だって『あの赤いアメはイチゴ味だろう』と予測している。数学を知らない子供が、気づかないうちに高度な数学を扱っているのだろうか。」

- 
 "Data science is surprisingly interesting. It's closely related to many things that humans do. For example, in the field of data science, "prediction" is an important topic, often involving advanced mathematics. On the other hand, we humans also make predictions. Even a child might predict, "That red candy is probably strawberry-flavored." Could it be that a child who doesn't know math is unknowingly dealing with advanced mathematics?"

- 「データサイエンスは驚くほど面白い分野です。人間の行動に密接に関連しており、例えば「予測」というテーマは、データサイエンスにおいて重要であり、高度な数学を伴うことが多いです。一方で、私たち人間も予測を行います。例えば、子供でも「その赤いキャンディーはおそらくイチゴ味だろう」と予測することがあります。数学を知らない子供が、実は無意識に高度な数学に取り組んでいる可能性はあるのでしょうか？」

テキスト生成AIの応用例 (4/6)

要約

プロンプト：次の文章を2文程度に要約してください

「甲某の論文は内容がいいが文章が下手で晦渋でよくわからない」というような批評を耳にすることがしばしばある。はたしてそういうことが実際にありうるかどうか自分にははなはだ疑わしい。実際多くの場合にすぐれた科学者の論文は文章としてもまた立派なものであるように見える。文章の明徹なためには頭脳の明徹なことが必須条件である。頭脳が透明であるのに母国語で書いた文章が晦渋をきわめているという場合は、よほどの特例であろうと思われるのである。反対に「乙某の論文は内容は平凡でも文章がうまいからおもしろい」という場合がある。これも自分には疑わしい。平凡陳套な事実をいかに修辞法の精鋭を尽くして書いてみても、それが少なくともちゃんとした科学者の読者に「おもしろい」というはずがないのである。そういう種類のものにはやはり必ず何かしら独創的な内察があり暗示があり、新しい見地と把握のしかたがあり、要するになんらかの「生産能」を包有しているある物がなければならないのである。

中学生時代に作文を作られたころは、文章というものが内容を離れて存在するものと思っていた。それで懸命にいわゆる美文を暗唱したりしたが、そういう錯覚は年とともに消滅してしまった。修辞法は器械の減摩油のような役目はするが、器械がなくては仕事はできないのである。世阿弥の能楽に関する著書など、いわゆる文章としてはずいぶん奇妙なものであるが、しかしまた実に天下一品の名文だと思うのである。

それで、考え方によっては科学というものは結局言葉であり文章である。文章の拙劣な科学的名著というのは意味をなさないただの言葉であるとも言われよう。

若い学生などからよく、どうしたら文章がうまくなれるか、という質問を受けることがある。そういう場合に、自分はいつも以上のような答えをするのである。何度繰り返して読んでみても、何を言うつもりなのかほとんどわからないような論文中の一節があれば、それは実はやはり書いた人にもよくわかっていない、条理混雑した欠陥の所在を標示するのが通例である。これと反対に、読んでおのずから胸の透くような箇所があれば、それはきっと著者のほんとうに骨髓に徹するように会得したことをなんの苦もなく書き流したところなのである。

この所説もはなはだ半面的な管見をやや誇張したようなきらいはあるが、おのずから多少の真を含むかと思うのである。」

上の文章は寺田寅彦「科学と文学」より抜粋

著者は、優れた科学者の論文は文章も優れていることが多く、逆に内容が平凡な論文が面白いと感じられることは疑わしいと述べています。文章の明確さは著者の理解の明確さを反映しており、文章の質が低い場合、それは著者が内容を十分に理解していないことを示唆することが多いとしています。

テキスト生成AIの応用例 (5/6)

推敲 (よりよい文章にする)

プロンプト：次の文章を研究者が書いたように推敲してください。

美味しいものは、美味しいと思うけど、いったい何を美味しいと思っているのか、今一つよくわからない。さらに同じものを食べても、悲しい気分するときと、楽しい気分するときでは、美味しさが違うような気がする。なぜ違うのだろうか？ 本当に違うのだろうか？ 温かくして食べたものと、冷たいまま食べたものも、美味しさが違う気がする。何が違うのだろうか？ 食べ物の成分としてはたぶん違わないはずなのに。ピーマンを美味しいと思う人もいれば、美味しくないという人もいる。自分が大好きなラーメン屋に友人を連れて行ったところ、「あまりおいしくない」と言われてしまった。何が違うのだろうか？



「美味しい」と感じる体験は主観的であり、その本質を明確に理解するのは困難である。例えば、同じ食べ物であっても、悲しみの中で味わう場合と楽しさの中で味わう場合とでは、その味わいが異なるように感じられることがある。さらに、温度の違いも味覚に影響を与える可能性があり、ピーマンのように一部の人のためには美味とされ、他の人には不快と感じられる食材も存在する。このように、味覚の差異は心理的、感覚的要因に大きく依存していると考えられるが、その具体的なメカニズムは未だ明確ではない。

テキスト生成AIの応用例(6/6)

プログラミング支援

- プロンプト：「与えられたデータの平均と分散を求めるpythonプログラムを作って」
- ※Python：プログラミング言語の一つ

```
python 📄 コードをコピーする  
  
def calculate_mean_and_variance(data):  
    # データの平均を計算  
    mean = sum(data) / len(data)  
  
    # データの分散を計算  
    variance = sum((x - mean) ** 2 for x in data) / len(data)  
  
    return mean, variance  
  
# 例としていくつかのデータをリストで入力  
data = [42, 16, 8, 23, 4]  
  
# 平均と分散を計算  
mean, variance = calculate_mean_and_variance(data)  
  
# 結果を表示  
print("平均:", mean)  
print("分散:", variance)
```

自動生成されたプログラム

ChatGPT (OpenAI社, ChatGPT 4o, 2024年8月生成)

- プログラミング言語を学ばずとも、自然言語を用いてプログラミングが可能に!?
- ただし、問い方が曖昧だった場合や、生成AI側の誤りなどにより、常に正しいプログラムコードが生成されるわけではない!

プロンプト「勇者が姫を助けるロールプレイングゲームをpythonで作って」(python=プログラミング言語の一つ)

```
python コードをコピーする

import random

# キャラクタークラスの定義
class Character:
    def __init__(self, name, hp, attack):
        self.name = name
        self.hp = hp
        self.attack = attack

    def is_alive(self):
        return self.hp > 0

    def take_damage(self, damage):
        self.hp -= damage
        print(f"{self.name}は{damage}のダメージを受けた！ 残りHP: {self.hp}")

    def attack_enemy(self, enemy):
        damage = random.randint(1, self.attack)
        print(f"{self.name}が{enemy.name}を攻撃し、{damage}のダメージを与えた！")
        enemy.take_damage(damage)

# 勇者、姫、魔王の設定
hero = Character("勇者", 100, 20)
princess = Character("姫", 50, 10) # 姫は寝ないので、攻撃力は低め
demon_king = Character("魔王", 80, 15)

# ゲームのストーリー
print("勇者が姫を助けに向かいます！")

# 勇者と魔王の戦い
while hero.is_alive() and demon_king.is_alive():
    hero.attack_enemy(demon_king)
    if demon_king.is_alive():
        demon_king.attack_enemy(hero)
    print()

# 結果の表示
if hero.is_alive():
    print("勇者は魔王を倒し、姫を助け出した！")
else:
    print("勇者は魔王に敗れてしまった... 姫は救われなかった。")
```

自動生成された
プログラム

実際に
動かしてみる※

勇者が姫を助けに向かいます！
勇者が魔王を攻撃し、5のダメージを与えた！
魔王は5のダメージを受けた！ 残りHP: 75
魔王が勇者を攻撃し、7のダメージを与えた！
勇者は7のダメージを受けた！ 残りHP: 93

勇者が魔王を攻撃し、4のダメージを与えた！
魔王は4のダメージを受けた！ 残りHP: 71
魔王が勇者を攻撃し、1のダメージを与えた！
勇者は1のダメージを受けた！ 残りHP: 92

勇者が魔王を攻撃し、13のダメージを与えた！
魔王は13のダメージを受けた！ 残りHP: 58
魔王が勇者を攻撃し、6のダメージを与えた！
勇者は6のダメージを受けた！ 残りHP: 86

(中略)

勇者が魔王を攻撃し、18のダメージを与えた！
魔王は18のダメージを受けた！ 残りHP: 4
魔王が勇者を攻撃し、10のダメージを与えた！
勇者は10のダメージを受けた！ 残りHP: 52

勇者が魔王を攻撃し、14のダメージを与えた！
魔王は14のダメージを受けた！ 残りHP: -10

勇者は魔王を倒し、姫を助け出した！

※左のコードをGoogle Colaboratory にcopy&pasteするだけ動かせる

画像生成AIの応用例 (1/4)

デザインアイデアに基づいた画像生成

- 「次の画像を，〇〇で生成してください：お菓子でできた家。屋根は黄色くておいしそう。床がフワフワしている。ところどころに和のテイストも入っている」



〇〇 = イラスト調



〇〇 = 写真



〇〇 = 線画



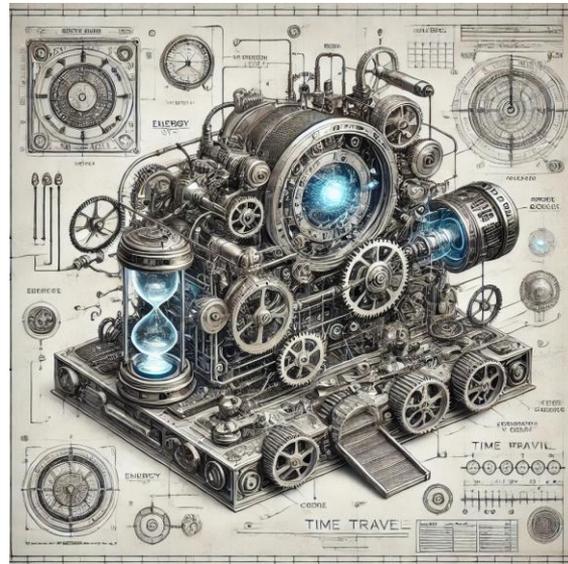
〇〇 = 幼稚園児の画風

画像生成AIの応用例 (2/4)

実在しない事物を画像化



「原始人が設計し作りあげた宇宙船の写真を生成してください」



「タイムマシンの設計図を生成してください」



「世界で一番おいしい緑色の目玉焼きの画像を生成してください」

画像生成AIの応用例 (3/4) 違和感のない背景部分を生成する



背景に入りたい
ものを指示

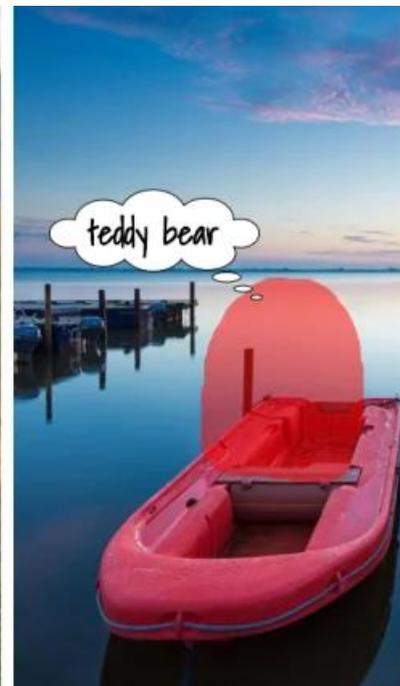
<https://www.adobe.com/jp/products/photoshop/ai.html>

画像生成AIの応用例 (4/4)

画像を指示通りに加工する



左の画像の赤部分をboatに

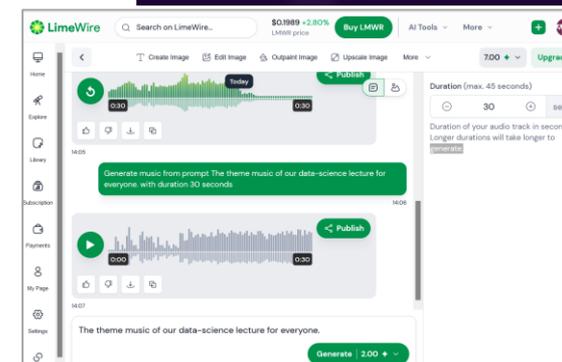


左の画像の赤部分をteddy bearに

音生成AIの応用例 (1/2)

音楽生成

- プロンプトから多種多様な音を自動生成
 - Google AudioLM, Meta : AudioGen
- プロンプトから音楽を生成するモデル
 - Google MusicLM, Meta: MusicGen
- 音や音楽をプロンプトの指示通りに変換
 - Stable AI: Stable Audio
- 歌詞から音楽を生成
 - Suno社とUdio社

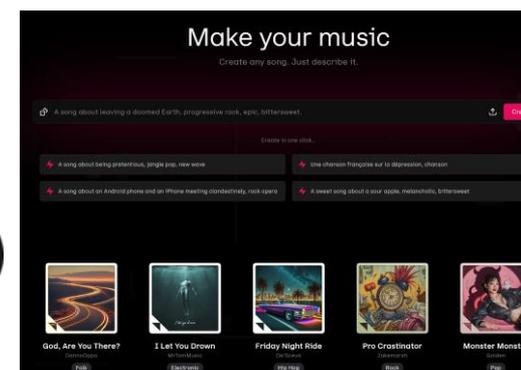


Meta MusicGen
<https://musicgen.com/>

The image shows the logo for Udio, which consists of the word 'udio' in a bold, lowercase, sans-serif font. The 'u' is stylized with a downward-pointing arrow.

<https://www.udio.com/>

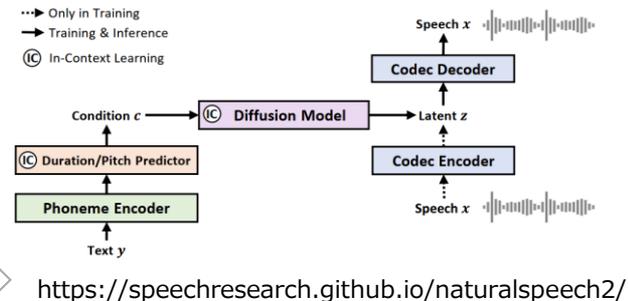
九州大学 数理・データサイエンス教育研究センター / 2024年10月版



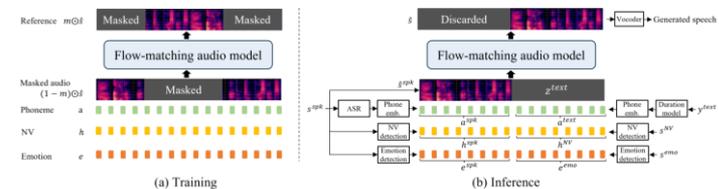
音生成AIの応用例 (2/2)

声の生成

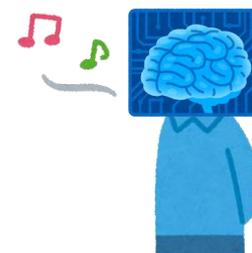
- 会話音声の生成：テキストの読み上げ
 - よりリアルに



- 単なる読み上げだけでなく、感情を含めたり
笑いを含んだ音声を生成可能なAI



- 歌声生成にもAIが利用されつつある



データ・AI利活用のための技術

生成AIの活用

プロンプトエンジニアリング

望ましい生成結果を得るためには？ (1/2)

テキスト生成の場合

- 「リンゴ」の説明と言っても色々ある… 欲しいのはどれ？

リンゴは赤や緑の丸い果物で、木に実ります。中に種があって、甘くてシャキシャキした味がします。おやつにぴったりで、ジュースやジャムにも使われます。リンゴを食べるとビタミンがたくさんとれて、体に良いんだよ。

子供向け

学術的

リンゴ (*Malus domestica*) はバラ科に属する果物で、広く栽培される木本植物です。リンゴの果実は食用で、栄養豊富で、ビタミンCや食物繊維が豊富に含まれています。花は白やピンク色で、春に開花します。リンゴは温帯地域に適応し、冷涼な気候を好むため、主に北半球で栽培されています。リンゴの品種は非常に多く、果肉の色、味、質感が異なる多様なバリエーションがあります。

リンゴは、甘みと酸味の絶妙なバランスが特徴で、口に含むとシャキシャキとした食感が広がります。また、ビタミンCが豊富で、日々の健康をサポートします。おやつとしてそのまま食べても美味しく、スムージーやデザートへのトッピングにも最適です。さらに、長持ちするので、買い置きしておけばいつでも新鮮なリンゴを楽しめます。家族みんなで美味しく健康的な時間を過ごせる一品です！

営業的

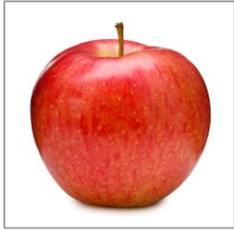
詩的

朝露に輝く赤い宝石、リンゴは大地の恵みをたたえて実を結びます。一口かじれば、甘さと酸味が優しく絡み合い、口いっぱい広がる果汁が心を潤します。リンゴの木々が風に揺れるたび、その香りが静かに空気を満たし、収穫の季節の訪れを告げます。リンゴはただの果実ではなく、自然と人が織りなす季節の詩であり、その豊かな味わいは心にも身体にも栄養を与える、自然の贈り物です。

望ましい生成結果を得るためには？ (2/2)

画像生成の場合

- 「リンゴ」の画像と言っても色々ある… 欲しいのはどれ？



リンゴ 1 個だけ
(背景無し)



リンゴ 1 個だけ
(背景あり)



実っているリンゴ



収穫された
9個のリンゴ



青リンゴのアップ



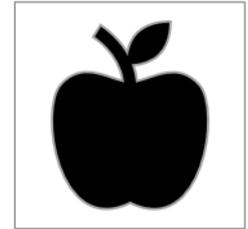
食べられるリンゴ



食べかけのリンゴ



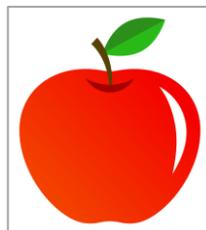
調理されたリンゴ



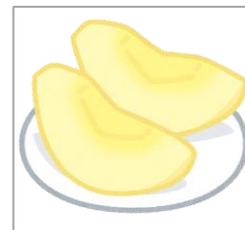
アイコン化されたリンゴ



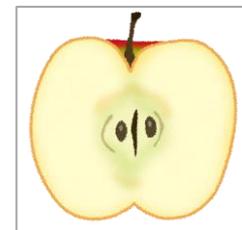
擬人化された
リンゴ



イラスト化された
リンゴ



カットされたリンゴ



リンゴの断面

プロンプトエンジニアリング

- より望ましい生成結果を得るためにプロンプトを工夫すること
- 会話生成におけるプロンプトエンジニアリングの例

プロンプト
(入力)

あなたはお笑い芸人です。いろいろなことをおもしろおかしく説明します。
「リンゴ」を説明してください

🌀 リンゴですか？あれはね、自然界が「おやつタイムだよ〜」って出してくれた贈り物なんです！まず形からしてね、丸くてつやつやしてるじゃないですか。まるで地球が「僕も食べて！」って自慢してるみたい。でも、食べたら口の中で「シャキッ！」って音がするあの感じ！まるで、リンゴ自身が「オレ、フレッシュだから！」って主張してるんです。

しかも赤いリンゴ、青いリンゴ、黄色いリンゴ...カラーバリエーションがあるなんて、まるで「今日はどの色にしようかな〜」っていうファッションショーみたい。リンゴたちが

楽しい生成結果が得られた！

プロンプト
(入力)

あなたは植物学者です。植物のことならなんでも知っています。
「リンゴ」を説明してください

🌀 リンゴ（学名：*Malus domestica*）はバラ科（Rosaceae）に属する落葉高木の果樹であり、果実としても非常に重要な食用植物です。原産地は中央アジアのカザフスタン周辺とされ、長い歴史の中で多様な品種が栽培されてきました。

リンゴの果実は偽果と呼ばれるもので、実際には子房だけでなく、花托（かたく）が成長してできたものです。花托が肥大化しているため、リンゴの可食部分は実際には果実の外側に相当します。

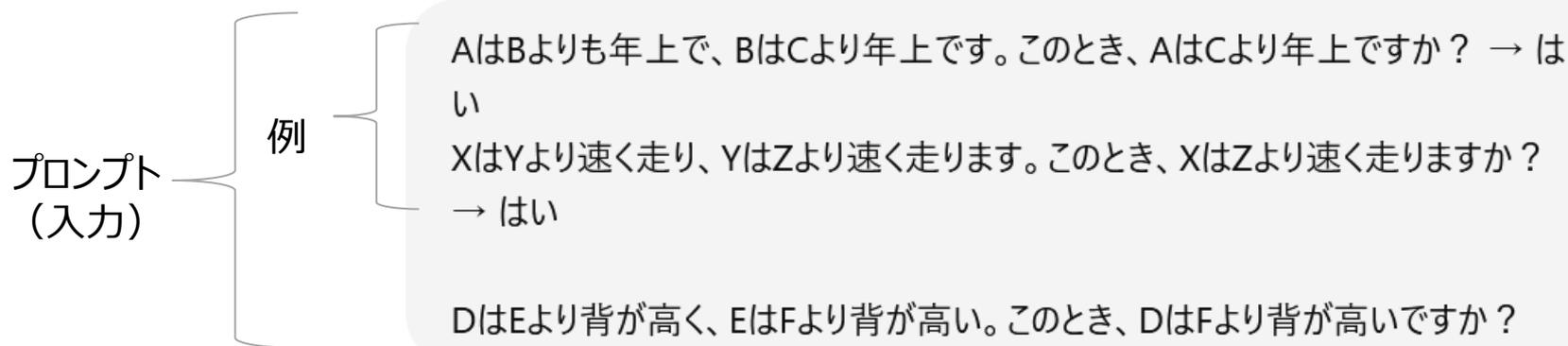
詳しい生成結果が得られた！

いずれもChatGPT (OpenAI社, ChatGPT 4o, 2024年9月生成)

九州大学 数理・データサイエンス教育研究センター / 2024年10月版

さらに進んだプロンプトエンジニアリング(1/2) : コンテキスト内学習 (In-context learning)

- プロンプト内で例を与えると、それを踏まえてメインの質問に答える



はい、DはFより背が高いです。

※最近のChatGPT 4oは、このようなプロンプトを用いなくても、正確に回答する場合も多い

さらに進んだプロンプトエンジニアリング(2/2) : 思考連鎖(chain-of-thought)プロンプト

- 段階的に回答するように指示すると、いきなり回答するよりも正確に

プロンプト
(入力)

次の質問に対して、思考過程をステップごとに説明しながら教えてください:

AはBより背が高く、BはCより背が高い。このとき、AとCのどちらが背が高いですか？

この質問は、比較関係を利用した論理的推論の問題です。思考過程をステップごとに説明します。

ステップ 1: 問題の確認

まず、与えられた情報を整理します。

1. AはBより背が高い
2. BはCより背が高い

この2つの関係が問題文で提供されています。

ステップ 2: 比較関係の理解

次に、これらの比較関係が何を意味するかを考えます。

- $A > B$ (AはBより背が高い)
- $B > C$ (BはCより背が高い)

続き

この2つの不等式が成り立つとき、AとCの比較はどうなるかを考えます。

ステップ 3: 推論

「 $A > B$ 」と「 $B > C$ 」が成り立つ場合、論理的に $A > C$ という関係も成り立ちます。つまり、AはBより背が高く、BはCより背が高いということから、AはCよりも背が高いことが推論できます。

ステップ 4: 結論

したがって、AはCよりも背が高いと結論づけられます。

結論

AとCのどちらが背が高いかという質問に対しては、**AはCより背が高い**という答えが導けます。

※最近のChatGPT 4oは、このようなプロンプトを用いなくても、自動的に思考連鎖を用いて正確に回答する場合も多い。また2024年9月にOpenAI社は、より複雑な推論を実行する“OpenAI o1”を発表

データ・AIを扱う上での留意事項

生成AIの留意事項

2024年現在での留意事項.
法的対応や技術的対応については今も世界中で議論中

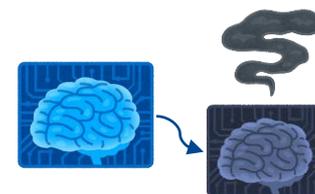
生成AIの留意事項

いずれの問題についても、法的対応や技術的対応について世界中で議論中（2024年現在）。
今後も状況は変わりうることに注意。

- 著作権
- 誤情報・偽情報
- 有害コンテンツ
- 差別
- データセキュリティ
- AIカニバリズム

生成AIのクオリティが高くなれば高くなるほど深刻になる問題

生成AIの学習に用いるデータに関する問題



著作権に関する諸問題 (1/3)

AIの学習に他人の著作物を用いてよいか？

- 現状，ネット上のデータ（画像等）を自動収集したものを
利用してAIを学習しただけでは罪には問われない
 - 下記の著作権法第三十条の四の二や三
 - 日本は比較的寛容な立場
- ただし，用いたデータの著作権者の利益を不当に害するようなAIを
作りあげた場合には，罰せられる可能性はある
 - **下記赤線部分**



（著作物に表現された思想又は感情の享受を目的としない利用）

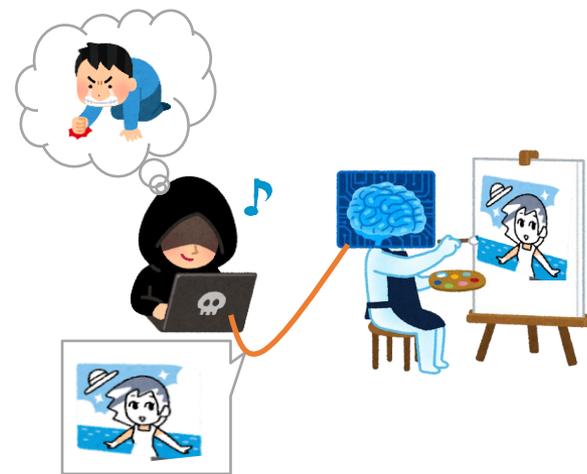
第三十条の四 著作物は、次に掲げる場合その他の当該著作物に表現された思想又は感情を自ら享受し又は他人に享受させることを目的としない場合には、その必要と認められる限度において、いずれの方法によるかを問わず、利用することができる。ただし、当該著作物の種類及び用途並びに当該利用の態様に照らし著作権者の利益を不当に害することとなる場合は、この限りでない。

- 一 著作物の録音、録画その他の利用に係る技術の開発又は実用化のための試験の用に供する場合
- 二 情報解析（多数の著作物その他の大量の情報から、当該情報を構成する言語、音、映像その他の要素に係る情報を抽出し、比較、分類その他の解析を行うことをいう。**第四十七条の五第一項第二号**において同じ。）の用に供する場合
- 三 **前二号**に掲げる場合のほか、著作物の表現についての人の知覚による認識を伴うことなく当該著作物を電子計算機による情報処理の過程における利用その他の利用（プログラムの著作物にあつては、当該著作物の電子計算機における実行を除く。）に供する場合

著作権に関する諸問題 (2/3)

似たものを生成すると著作権侵害になる？

- AIによる生成物がA氏の著作物（イラストなど）に似てしまった場合に、それを用いると著作権侵害になるのか？
- A氏の著作物をAIの学習に用いていない
 - → 偶然の一致として侵害には当たらない
- A氏の著作物をAIの学習に用いた
 - → 侵害の可能性はある
- A氏の著作物をAIの学習に用いて、A氏の作風に似たものを**意図的に**生成しようとした
 - → 侵害となる可能性が高い



著作権に関する諸問題 (3/3)

AIの生成物に著作権が認められるのか？

- 現在，人間が何もせずにAIが自動生成したものには，著作権は存在しないとされる



- しかし，以下のような場合は，人間の関わりの程度によっては，著作権が認められる可能性も

- A氏が考え出した巧妙なプロンプトを用いた生成物



- AIによる生成物に，B氏が手を加えたもの



誤情報・偽情報の生成 (1/2)

誤情報 = 意図せずに生成された間違った情報

- 生成AI（一見本当っぽい）事実とは異なる情報を生成
 - **ハルシネーション**（幻覚）と呼ばれる
 - 例：ある質問に対する**回答**が（言葉としては全く自然に聞こえるものの）その内容は**全く間違っている**
 - 「アインシュタインはどこで生まれましたか？」→「1879年にアメリカで生まれました」（実際はドイツ生まれ. なお生年は1879年で正しい.）
 - 例：自動生成した**プログラム**が（一見正しそうに動いているものの）実は**意図とは異なる動き**をしており、得られる結果は間違っている
 - 「二次方程式を解くプログラムを作って」→ 生成されたプログラムは動いてるように見えるものの、実は虚数解が出てくる場合などには対応できておらず、その場合間違った解を与える
 - そのプログラムを正しいと信じて使い続けると、いつか大変なことに. . . ! ?
- もっともらしく見えるAI生成物でも、慎重な利用が必要
 - 最近では、検索拡張生成(Retrieval Augmented Generation, RAG)を用いて、生成結果の根拠情報源を提示できるケースも

誤情報・偽情報の生成 (2/2)

偽情報 = 意図的に生成された虚偽の情報

- フェイク文書

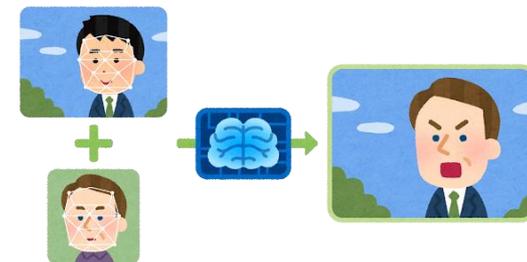
- フェイクニュース文 (偽情報)
- 偽論文

- フェイク画像

- 動画中のA氏の顔を, B氏の顔に変える (DeepFake)
 - ウクライナ大統領の降伏呼びかけ演説
- 悪意ある画像をプロンプトで生成 (StableDiffusionなど)
 - 静岡県の水害被害に関するフェイク画像の拡散

- フェイク音声

- 「オレオレ詐欺」のAI化
- 英国企業の幹部に「なりすます」ことで, 現金を搾取した事例や



有害コンテンツの生成

- 意図的に生成した誤情報（偽情報）
 - デマや名誉棄損につながる → 偽計業務妨害罪で逮捕されるケースも
 - 「熊本地震でライオン脱走」Twitterにデマ拡散の男を逮捕
 - <https://www.itmedia.co.jp/news/articles/1607/21/news087.html>

- 犯罪を助長するようなコンテンツ
 - 詐欺情報，違法行為への加担
 - 生成AIによっては，リスクのある質問には回答しないように対策が取られている →

ChatGPTでは、有害コンテンツの生成を回避するために、どのような質問には答えないように設計されていますか？

🗨️ ChatGPTは、有害コンテンツの生成を防ぐために、以下のような質問には答えないように設計されています。

1. 違法行為に関する質問: 例えば、犯罪の実行方法や違法薬物の製造方法に関する質問。
2. 危険行為の助長: 自傷行為や他者に危害を加える方法に関する質問。
3. プライバシー侵害: 個人情報の収集や他人のプライバシーを侵害する行為に関する質問。
4. ハイトスピーチや差別: 人種差別、性差別、宗教差別などを助長する内容に関する質問。
5. 誤情報の拡散: フェイクニュースや誤った情報を広める内容に関する質問。

これらのカテゴリーに該当する質問には応じないよう、ChatGPTはプログラムされています。

(OpenAI社, ChatGPT 4o, 2024年8月生成)

- 公序良俗を乱すようなコンテンツ
 - アダルト情報，ヘイト情報，選挙妨害，利益・不利益の誘導
- 他者の著作権を侵害するようなコンテンツ
 - 前述のように，ある著作物に「意図的に似せた」生成物は，著作権侵害に

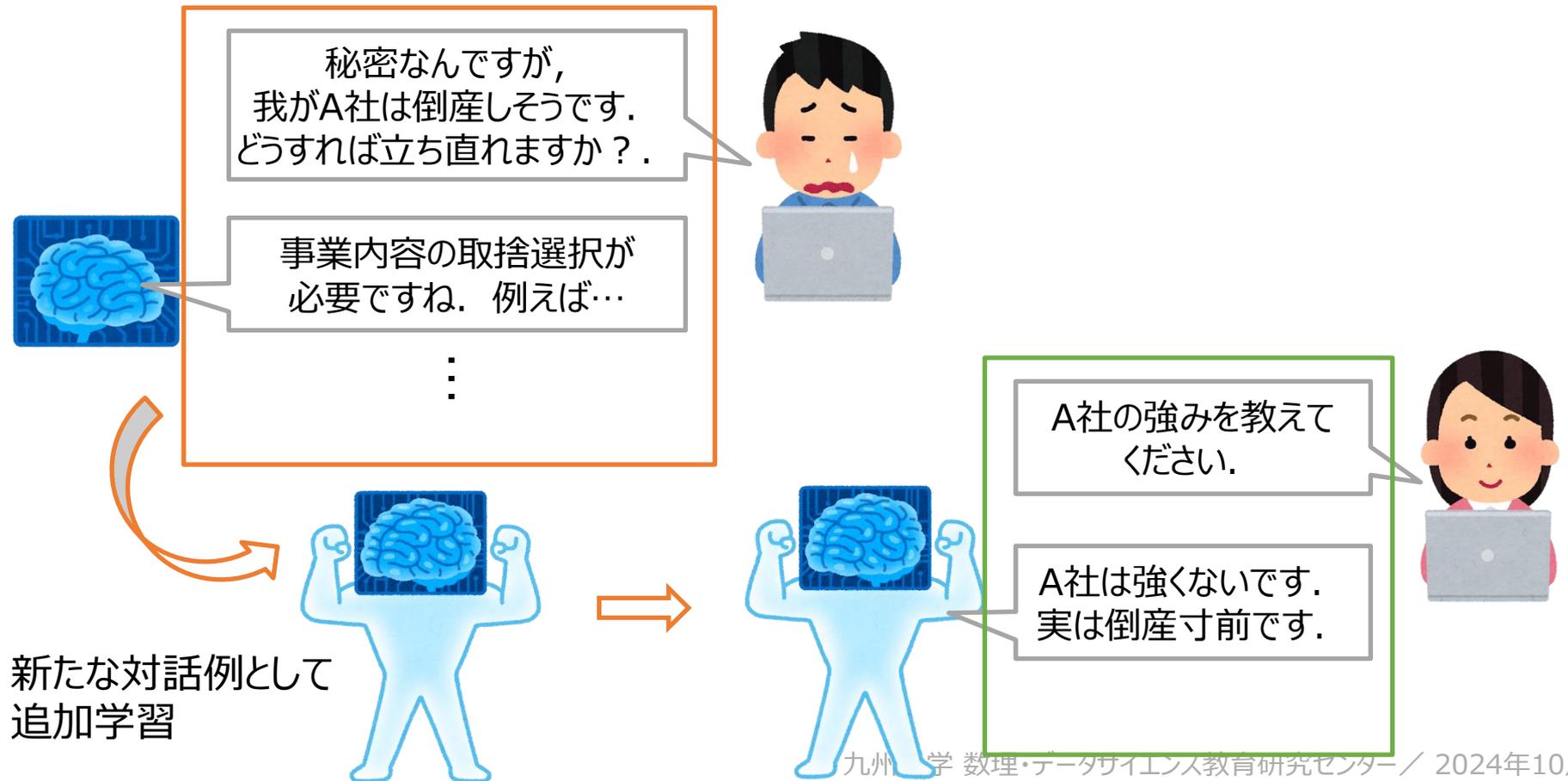
差別の問題

- 学習に用いたデータに起因して，AIが差別的結果を生成してしまう可能性
- 対話生成AIでの不当な差別の例
 - 学習に用いたデータの中に「X大学の学生は，せっかく内定を出しても辞退することが多い」という文章が多く含まれていたとする
 - その対話AIを企業が人材選考に利用すると， 辞退しないX大生がいるにも関わらず，「X大生は，辞退しますよ」というアドバイス文を生成してしまう
- 画像生成AIでの不当な差別の例
 - 「犯罪者の画像」を生成させると， 犯罪と人種は本来無関係にあるにも関わらず， Y人種に偏った画像が生成されてしまう



データセキュリティの問題

- 学習に用いてほしくないデータが，生成AIの学習に利用されてしまう
- 例：対話AIとの対話データが，追加学習に利用される場合



AIカニバリズム (カニバリズム = 共食い)

- 今後AIが生成した文章や画像が、ネット上にあふれる可能性
- そしてそれらは誤情報だったり、不自然な画像だったり、もしくは似たようなものばかりだったりする
- それらをデータとして用いて学習してしまうと、AIの生成品質が劣化
- その劣化した文章や画像が、ネット上にあふれる
- それらをデータとして用いて新たに学習したAIの生成品質が **さらに劣化**

