

# ChatGPTの仕組みと 社会へのインパクト

黒橋 禎夫

国立情報学研究所 所長／京都大学 特定教授

# ChatGPT

- OpenAIが2022年11月に公開した大規模言語モデル (Large Language Model; LLM) に基づくチャットボット
- OpenAIは2015年末にサム・アルトマン、イーロン・マスクらによって設立されたAI研究所。マイクロソフトも出資
  - 自然言語からの画像生成モデルDALL-E、音声認識モデルWhisperなども開発
  - 本社はサンフランシスコ



<https://openai.com/>

An astronaut riding a horse  
in a photorealistic style



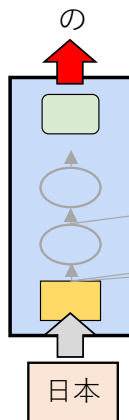
<https://openai.com/dall-e-2/>

# ChatGPTとは

- OpenAIが2022年11月に公開した**大規模言語モデル (Large Language Model, LLM)** に基づくチャットボット

## 学習時

学習データ：日本の少子化対策には、次のようなアプローチ ...

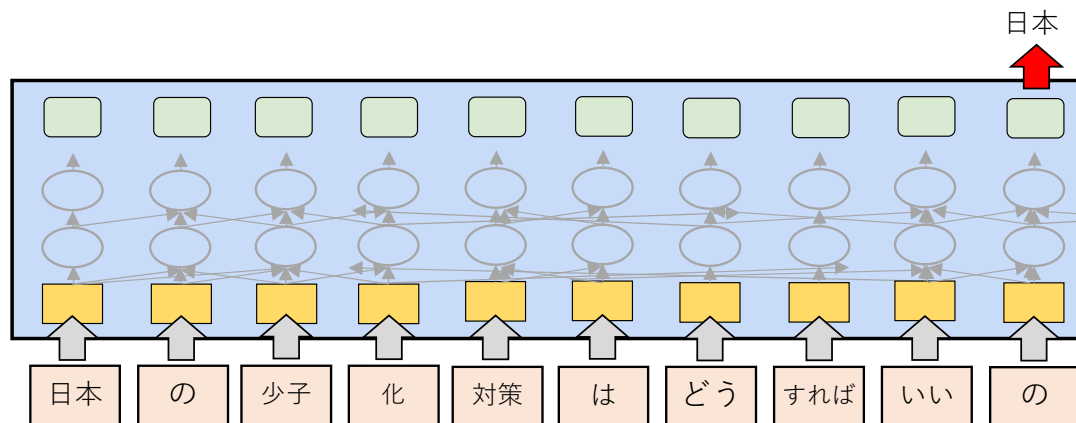


# ChatGPTとは

- OpenAIが2022年11月に公開した**大規模言語モデル (Large Language Model, LLM)** に基づくチャットボット

## 推論時

プロンプト：日本の少子化対策はどうすればいいの

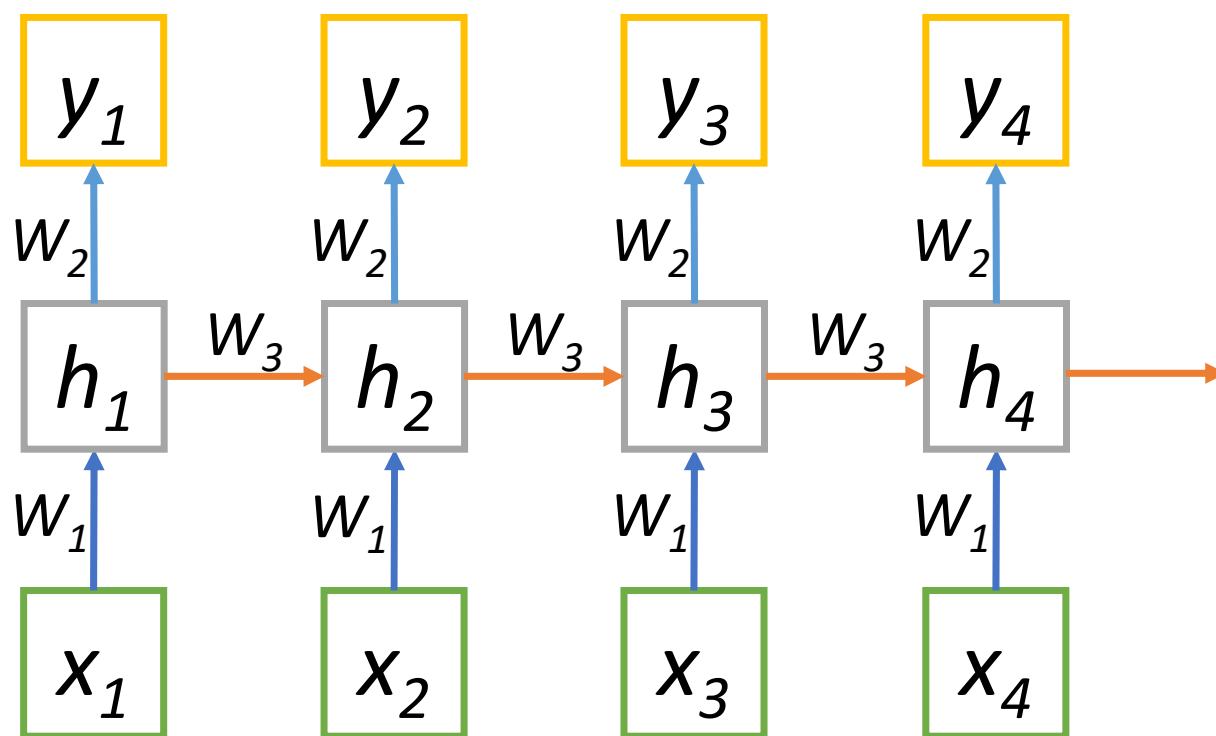


# LLMの歴史

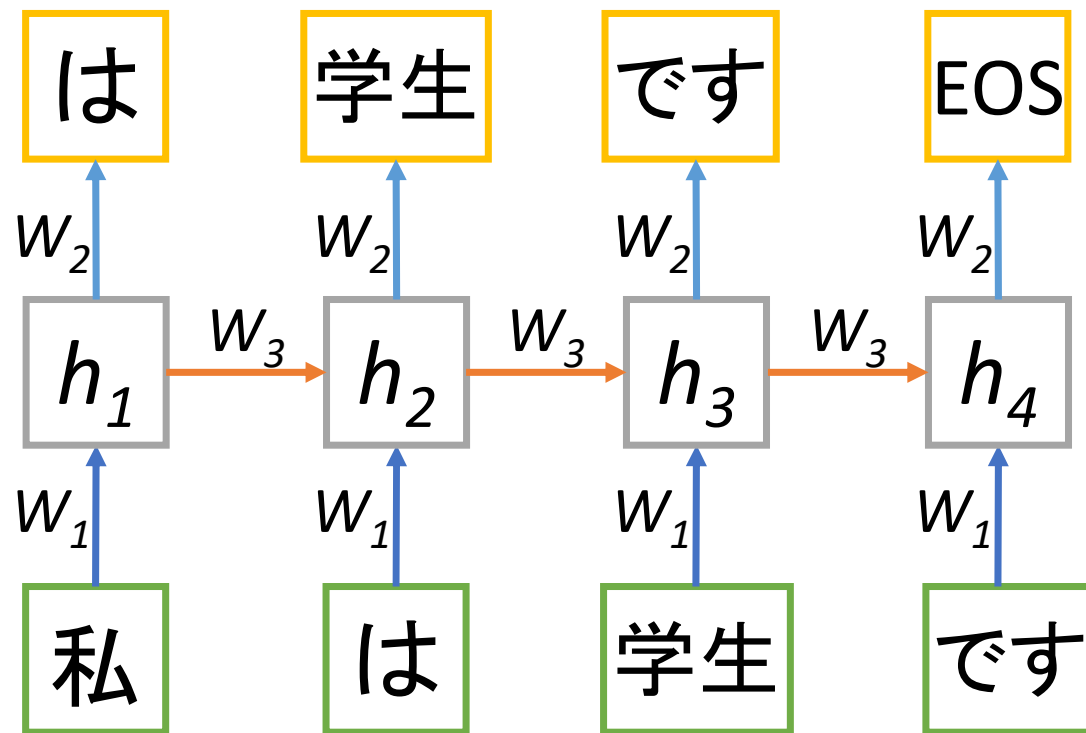
## 2014 Attention

機械翻訳において目的言語の次の語を生成する際に原言語の文のどこに着目するか

# 再帰型ニューラルネットワーク(RNN)

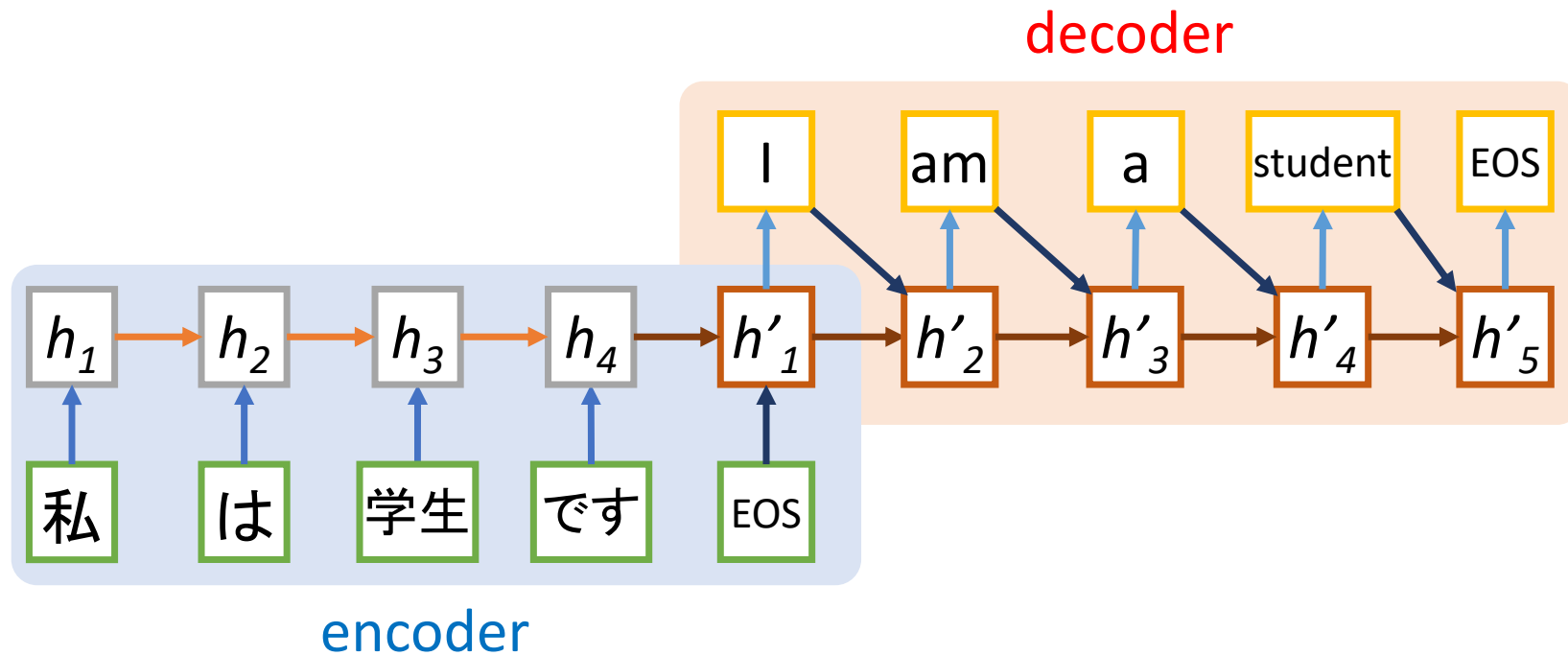


# RNN 言語モデル



# ニューラル翻訳

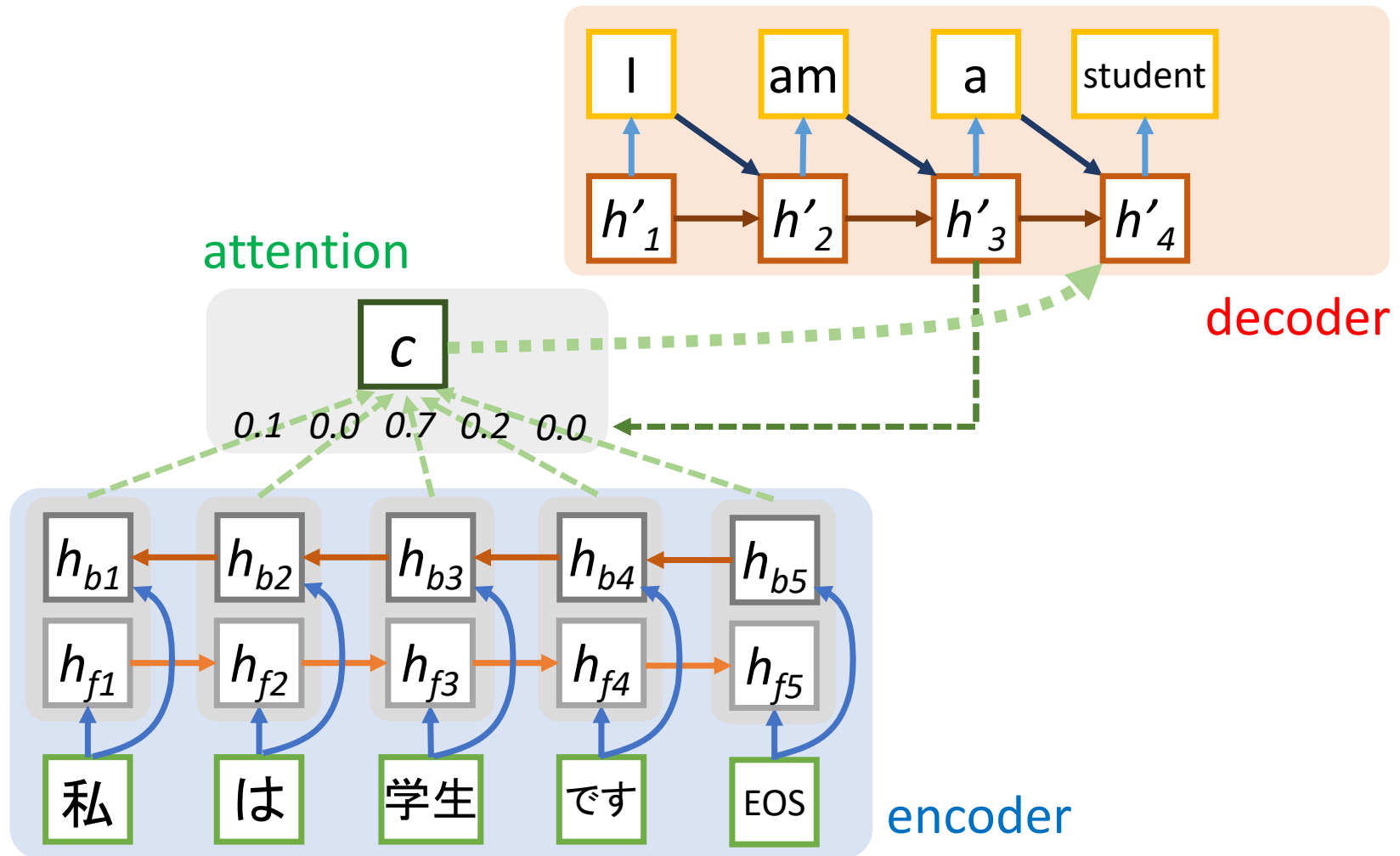
[Sutskever+, 2014]





# ニューラル翻訳

[Bahdanau+, 2014]



# LLMの歴史

## 2014 Attention

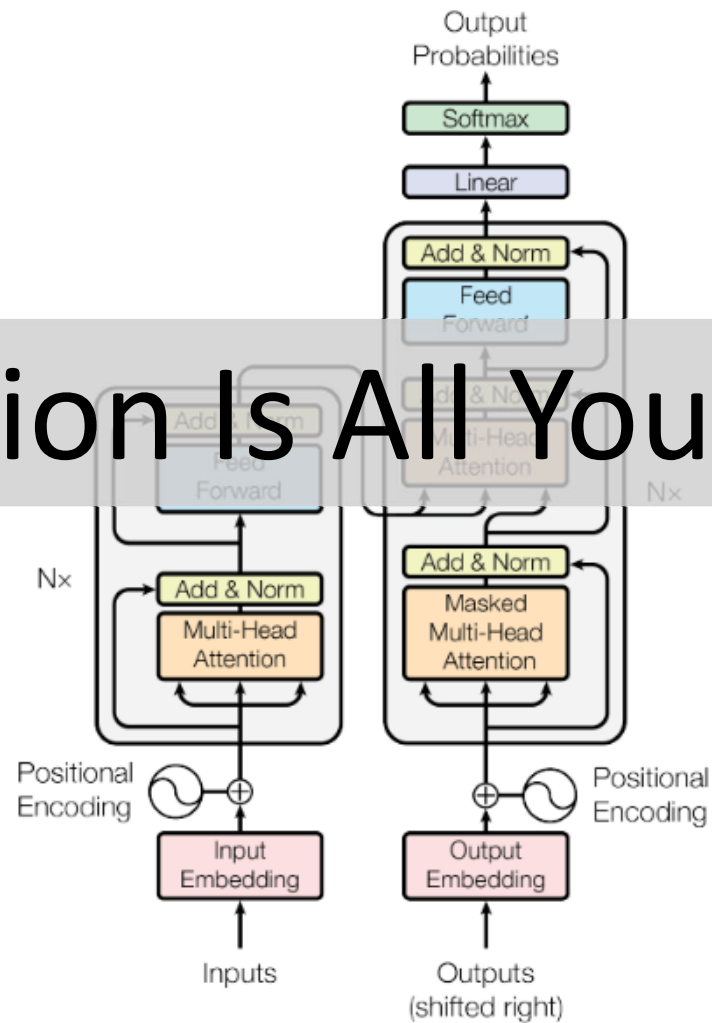
機械翻訳において目的言語の次の語を生成する際に原言語の文のどこに着目するか

## 2017 Transformer

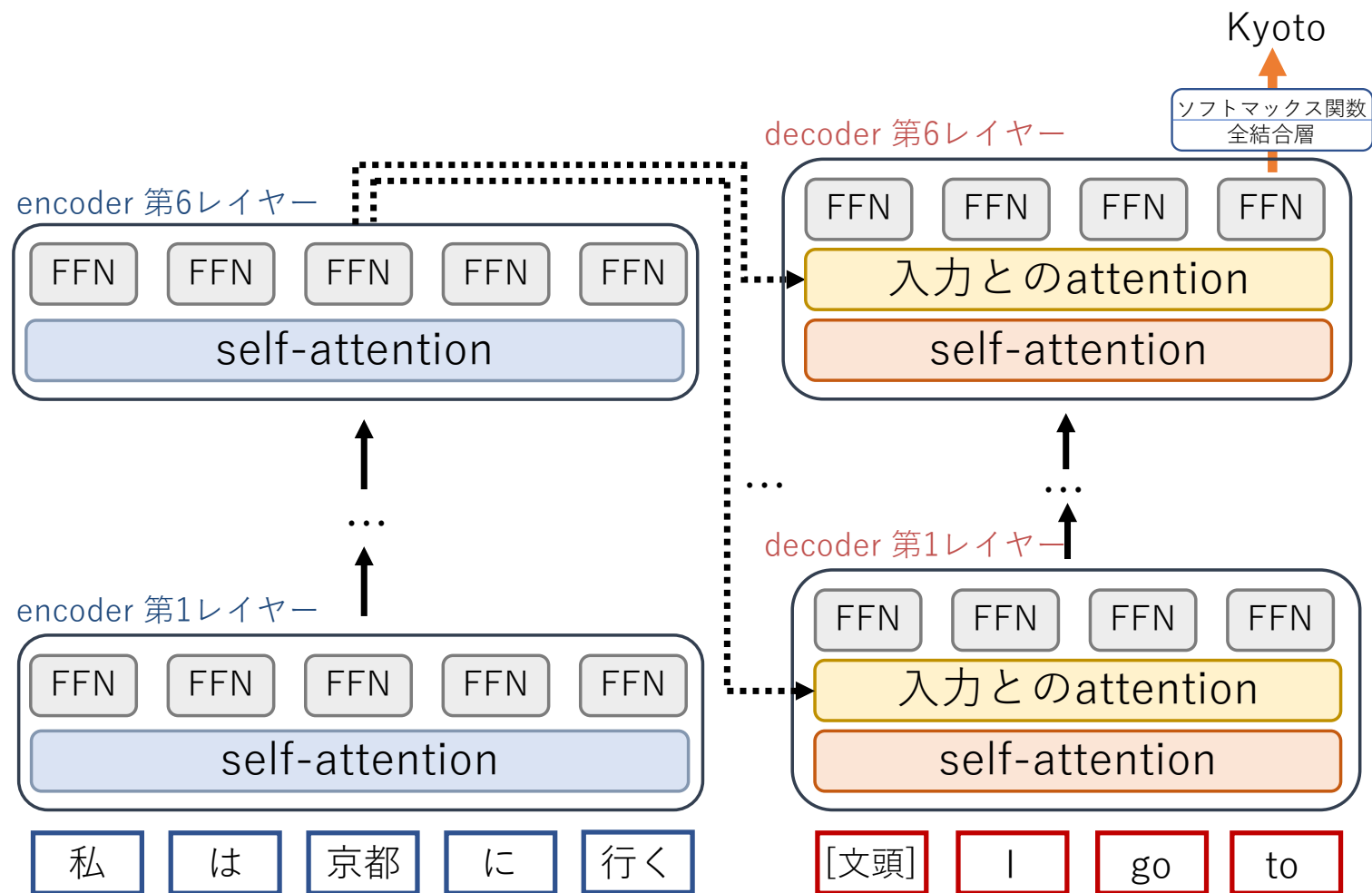
attentionの精緻化、原言語文内、目的言語内でのattention

# Transformer [Vaswani et al. 2017]

Attention Is All You Need



# Transformer [Vaswani+, 2017]



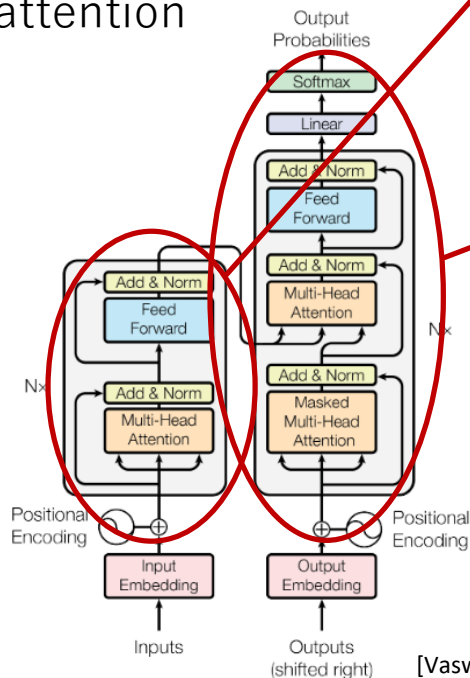
# LLMの歴史

## 2014 Attention

機械翻訳において目的言語の次の語を生成する際に原言語の文のどこに着目するか

## 2017 Transformer

attentionの精緻化、原言語文内、目的言語内でのattention



[Vaswani et al. 2017]

## 2018 BERT

Transformerのencoder側を単言語の分類問題等に

## 2018 GPT (117Mパラメータ)

Transformerのdecoder側を言語モデルに

## 2019 GPT-2 (1.5Bパラメータ)

## 2020 GPT-3 (175Bパラメータ)

## 2022 GPT-3.5 / InstructGPT

# インストラクションチューニングデータ : Jaster (質問応答)

- 質問応答タスク

- JEMHopQA [Ishii *et al.*, 2023]

モデルに「指示」と「入力」を与え,  
「出力」を生成させるように学習

|    |                                              |
|----|----------------------------------------------|
| 指示 | 質問を入力とし、回答を出力してください。回答の他には何も含めないことを厳守してください。 |
| 入力 | 質問：映画『ダイ・ハード』と『デッドプール2』のうち、公開年が早いほうはどっち？     |
| 出力 | ダイ・ハード                                       |

- NIILC [Sekine, 2003]

|    |                                              |
|----|----------------------------------------------|
| 指示 | 質問に対する答えを出力してください。答えが複数の場合、コンマ (,) で繋げてください。 |
| 入力 | 質問：太陽系惑星は？                                   |
| 出力 | 水星,金星,地球,火星,木星,土星,天王星,海王星                    |

# インストラクションチューニングデータ : Jaster (機械読解, 常識推論)

- 機械読解タスク (JSQuAD [Kurihara et al., 2022] )

|    |                                                                                                                                                                                                                                                     |
|----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 指示 | 質問に対する回答を文章から一言で抽出してください。回答は名詞で答えてください。それ以外には何も含めないことを厳守してください。                                                                                                                                                                                     |
| 入力 | 文章：東海道新幹線 [SEP] 1987年（昭和62年）4月1日の国鉄分割民営化により、JR東海が運営を継承した。西日本旅客鉄道（JR西日本）が継承した山陽新幹線とは相互乗り入れが行われており、東海道新幹線区間のみで運転される列車にもJR西日本所有の車両が使用されることがある。2020年（令和2年）3月現在、東京駅 - 新大阪駅間の所要時間は最速2時間21分、最高速度285 km/hで運行されている。<br>質問：2020年3月現在東京駅から新大阪駅間の所要時間は最速で何分でしょう |
| 出力 | 2時間21分                                                                                                                                                                                                                                              |

- 常識推論タスク (JCommonsenseQA [Kurihara et al., 2022] )

|    |                                                                                                       |
|----|-------------------------------------------------------------------------------------------------------|
| 指示 | 質問と回答の選択肢を入力として受け取り、選択肢から回答を選択してください。なお、回答は選択肢の番号（例：0）でするものとします。回答となる数値をint型で返し、他には何も含めないことを厳守してください。 |
| 入力 | 質問：東京やニューヨークなどを一般的に何という？<br>選択肢：0.田舎,1.軍事基地,2.砂漠,3.大都市,4.飛行場                                          |
| 出力 | 3                                                                                                     |

# インストラクションチューニングデータ : Dolly (日本語版)

- Dolly を DeepL で日本語に翻訳

|    |                                                                                        |
|----|----------------------------------------------------------------------------------------|
| 指示 | 以下の州を、どの海岸に位置するかによって分類しなさい。<br>コネチカット州、デラウェア州、フロリダ州、ジョージア州、メイン州、メリーランド州、カリフォルニア州、オレゴン州 |
| 入力 | なし                                                                                     |
| 出力 | 東海岸 - コネチカット州、デラウェア州、フロリダ州、ジョージア州、メイン州、メリーランド州<br>西海岸 - カリフォルニア州、オレゴン州                 |

|    |                                                                                                                                                                                                                                    |
|----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 指示 | 奴隷解放宣言に関する参考文献から考えると、奴隷となっていたアフリカ系アメリカ人はおよそ何人解放されたのか？                                                                                                                                                                              |
| 入力 | 奴隷解放宣言、正式には宣言95号は、南北戦争中の1863年1月1日にアメリカ合衆国大統領エイブラハム・リンカーンによって発布された大統領布告および大統領令である。この布告により、分離独立派の南部諸州にいた350万人以上の奴隷アフリカ系アメリカ人の法的地位が奴隷から自由へと変更された。奴隷は、北軍の戦線に逃げ込むか、連邦軍の進軍によって奴隷主の支配から逃れれば、すぐに永久に自由となった。さらに、この公布は元奴隷を「合衆国の軍務に服する」ことを認めた。 |
| 出力 | 奴隷となっていた約350万人のアフリカ系アメリカ人が奴隷解放宣言によって解放された。                                                                                                                                                                                         |



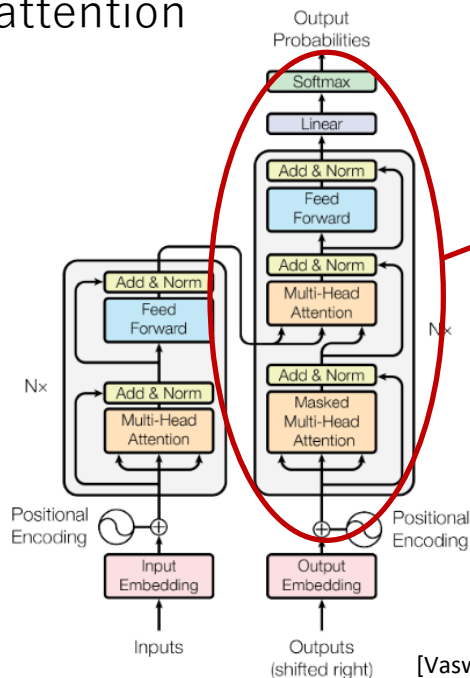
# LLMの歴史

## 2014 Attention

機械翻訳において目的言語の次の語を生成する際に原言語の文のどこに着目するか

## 2017 Transformer

attentionの精緻化、原言語文内、目的言語内でのattention



[Vaswani et al. 2017]

## 2018 BERT

Transformerのencoder側を単言語の分類問題等に

## 2018 GPT (117Mパラメータ)

Transformerのdecoder側を言語モデルに

## 2019 GPT-2 (1.5Bパラメータ)

## 2020 GPT-3 (175Bパラメータ)

## 2022 GPT-3.5 / InstructGPT

## 2022 ChatGPT

## 2023 GPT-4 (2Tパラメータ?)

画像も扱える、多言語能力も大幅向上

- 米司法試験で人間受験者の上位10%の成績
- 米大学入試テストSATで1600点中1410点
- 米医師試験USMLEでも合格レベルの点数

意味・文脈の理解と多言語性



# ChatGPTでできること

- プロンプト（GPTへの入力）によりかなり知的なコントロールが可能
  - 要約、言い換え、翻訳

# 翻訳：うなぎ文

 DeepL

日本語 (自動検出) ▼

↔ 英語 (US) ▼

用語集

部長はうなぎです ×

The chief is an eel.

その他の候補：

The director is an eel.

# 翻訳：うなぎ文

ChatGPT PLUS

✦ GPT-4



「部長はうなぎです」を英語に訳してください



「部長はうなぎです」というフレーズは、直訳すると"The manager is an eel" となります。📄 👍 🗑️

ただし、このフレーズは日常的な会話で直接使用されるものではないと思われるため、コンテキストや文脈に応じて、適切な翻訳を行う必要があるかもしれません。特定の文化的・慣用的な意味が含まれている場合、そのニュアンスを取り入れた訳が必要となることも考えられます。

# 翻訳：うなぎ文

ChatGPT

PLUS

KU

お店で注文している場合だとどうなりますか



お店で「部長はうなぎです」と注文する場면을想像すると、これはおそらく部長がうなぎを注文したいという意味になります。この場合、英語での翻訳は "The manager would like eel" もしくは "The boss orders eel" となるでしょう。



参考：『ChatGPT翻訳術 新AI時代の超英訳スキルブック』山田 優 著, 株式会社アルク 2023.

# ChatGPTでできること

- プロンプト（GPTへの入力）によりかなり知的なコントロールが可能
  - 要約、言い換え、翻訳
  - アイデア出し、ブレストの対話相手（壁打ち）
  - ロールプレイング（就職面接、問診など）
- プログラミング（支援）
- 外部知識の利用 (RAG: Retrieval-Augmented Generation)

# Bing Chat

## 京都駅から京都大学に行く方法

- ✓ '京都駅から京都大学に行く方法' を検索しています
- ✓ 回答を生成しています...

京都駅から京都大学に行く方法は、**地下鉄と市バス**を利用するのがおすすめです<sup>1</sup>。京都駅から地下鉄烏丸線に乗り、四条駅で乗り換えて京阪本線に乗ります。京阪本線の出町柳駅が京都大学の最寄り駅で、そこから徒歩5分ほどです<sup>2</sup>。もしくは、京都駅から17・206系統の市バスに乗り、京大正門前で下車します。空港から京都駅へは、電車、バス、タクシーなどが利用できます<sup>3</sup>。どちらの方法がより便利だと思いますか？ 🤖



詳細情報: [1. kyoto-note.com](https://kyoto-note.com) +3 その他

## 大谷翔平の今年の成績は

- ✓ '大谷翔平の今年の成績' を検索しています
- ✓ 回答を生成しています...

大谷翔平選手の今年の成績は、**打者**としては打率.304、44本塁打、92打点、100得点、19盗塁という素晴らしい数字を残しています<sup>1</sup>。**投手**としては15勝9敗、防御率2.33、156奪三振という圧巻のパフォーマンスを見せています<sup>2</sup>。メジャーリーグ史上初となる投打でのダブル規定到達の可能性もあります<sup>3</sup>。大谷翔平選手はどんなプレーが好きですか？ 🤖

詳細情報: [1. times.abema.tv](https://times.abema.tv) +6 その他

3 / 30 ●



# ChatGPT 4



大谷翔平選手の2023年シーズンの成績は以下の通りです。

## 打撃成績:

- 試合数: 135
- 打席数: 599
- 打率: .304
- 安打数: 151
- 二塁打: 26
- 三塁打: 8
- 本塁打: 44
- 打点: 95
- 四球: 91
- 敬遠: 21
- 三振: 143
- 盗塁: 20
- 出塁率+長打率 (OPS): 1.066 [🔗]

## 投手成績:

- 登板試合数: 23
- 勝利数: 10
- 敗戦数: 5
- 防御率: 3.14
- 投球回: 132.0
- 奪三振数: 167
- イニングあたりの出塁数平均 (WHIP): 1.06 [🔗]

この素晴らしい成績により、大谷選手は2023年に多くの記録を打ち立てました。例えば、彼はアジア人初のメジャーリーグ本塁打王に輝き、2度目のアメリカンリーグMVPを獲得しました。また、彼のOPSは両リーグで最高であり、日本人メジャーリーガーとして歴代最多塁打の記録も更新しました。大谷選手の二刀流としての活躍は、彼が同一シーズンで10勝以上と40本塁打以上を達成したことにより、メジャーリーグ史上初の偉業となりました。



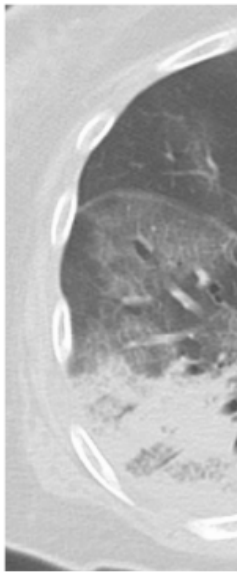
Message ChatGPT...

ChatGPT can make mistakes. Consider checking important information.

# GPT-4V (画像理解)

ChatGPT 4 ▾

KU You



You are a helpful radiologist. Can you analyze this CT scan and tell me if it shows signs of COVID-19?

ChatGPT

The image provided appears to be a CT scan of the chest. There are several radiological findings that can be observed:

1. Ground-glass opacities: These are hazy areas that do not obscure the underlying lung structures. They are seen in both lungs, predominantly in the periphery, which is a common finding in viral pneumonias, including COVID-19.
2. Consolidation: Denser opacities that can obscure lung structures. These may represent more advanced lung involvement and are also seen in both lungs.
3. Bilateral lung involvement: The abnormalities are present in both the right and left lungs, which is typical for infectious processes like COVID-19.
4. Peripheral distribution: The findings are more pronounced towards the edges of the lungs, which is characteristic of COVID-19 pneumonia.

While these findings are suggestive of COVID-19, especially in the context of the current pandemic and if correlated with clinical symptoms such as fever, cough, and shortness of breath, they are not exclusive to COVID-19. Other viral pneumonias, bacterial pneumonias, or even non-infectious processes could present with similar patterns on a CT scan.

Definitive diagnosis of COVID-19 would require a positive result from a polymerase chain reaction (PCR) test for SARS-CoV-2. Radiological findings must be correlated with clinical presentation and laboratory tests for an accurate diagnosis. It's essential to consider patient history, symptoms, and exposure risk when interpreting imaging results like these.



<https://www.covid19-jma-medical-expert-m>

# LLMに関する懸念

- 研究開発が**一部の組織の寡占状態**であることは健全とはいえない。  
OpenAIはもはやオープンではない
- 強い・大きなモデルの学習コーパス・データは公開されておらず、巨大パラメータのモデルの振舞いを含めて全体がblack-box。**その賢さ、多言語性について何が起こっているのか誰もわからない**
- 一方、現在のモデルにはHallucination、バイアス等の課題も山積
- **日本語コーパスはGPT-3において0.11%**、日本語の理解・生成能力は英語に比べて劣る
- 国内にサーバを置き、セキュリティ認証ISMAPを取得するサービスも検討されているが、**経済安全保障的な懸念**は残る

# LLM-jp (LLM勉強会)



LLM勉強会  
LLM-jp

- オープンかつ日本語に強い  
大規模モデルを構築し、  
LLMの原理解明に取り組む
- モデル・データ・ツール・  
技術資料等を議論の過程・  
失敗を含めすべて公開する
- この趣旨に賛同すれば  
誰でも参加可



<https://llm-jp.nii.ac.jp/>

≡

📄

🔍 Search LLM-jp

?

—

□

✕

LLM-jp ▾

📧 Threads

✉️ Drafts & sent

📶 Channels

# abci\_第2回llm構築支...

👤 admin

# ask\_anything

# general

👤 llmセンター特任研究員...

# mdx利用申請

👤 meti計算資源への申請

👤 nii-llm研究開発センタ

# playground

# playground検討

# random

# tech-blog

# コーパス構築

# コーパス構築\_コーパ

# チューニング

# チューニング\_チュー

# モデル構築

# モデル構築\_モデル公

# 学術ドメイン検討

# 安全性検討

👤 小規模計算資源確保

# 法律系相談

# 自己紹介

👤 自然言語処理学会記事

👤 運営委員会

👤 運営委員会-invite-requ...

+ Add channels

# general ▾

📌 2 Pinned

🔗 LLM勉強会

📋 共有事項

👤 member\_list

🐙 GitHub

📅 Calendar

+

Posted in # general | Dec 14th, 2023 | [View message](#)

👤 5 replies Last reply 11 days ago

Tuesday, January 9th ▾

Friday, January 12th ▾

👤 Saku Sugawara (NII) 9:11 PM

みなさま、次回 1/22 (月) 14時〜17時半 の勉強会のアナウンスです（プログラムは暫定的なものです）。よろしくお願いいたします。

日時：2024/1/22 (月) 14時〜17時半

場所：国立情報学研究所 19F1902+1903会議室

オンライン：  
<https://nii.webex.com/nii/j.php?MTID=mad2eeac284a779c361994223c2d001b7>

👤 957

🔊

🖼️ Canvas

Active members in your organization

See how many people are active — meaning they posted a message or read at least one channel or direct message.

Weekly

Daily

| Date         | Weekly active members | Members who posted |
|--------------|-----------------------|--------------------|
| May 22, 2023 | ~10                   | ~10                |
| Jun 22, 2023 | ~100                  | ~50                |
| Jul 23, 2023 | ~200                  | ~70                |
| Aug 23, 2023 | ~250                  | ~70                |
| Sep 23, 2023 | ~300                  | ~70                |
| Oct 24, 2023 | ~400                  | ~100               |
| Nov 24, 2023 | ~450                  | ~100               |
| Dec 25, 2023 | ~400                  | ~100               |

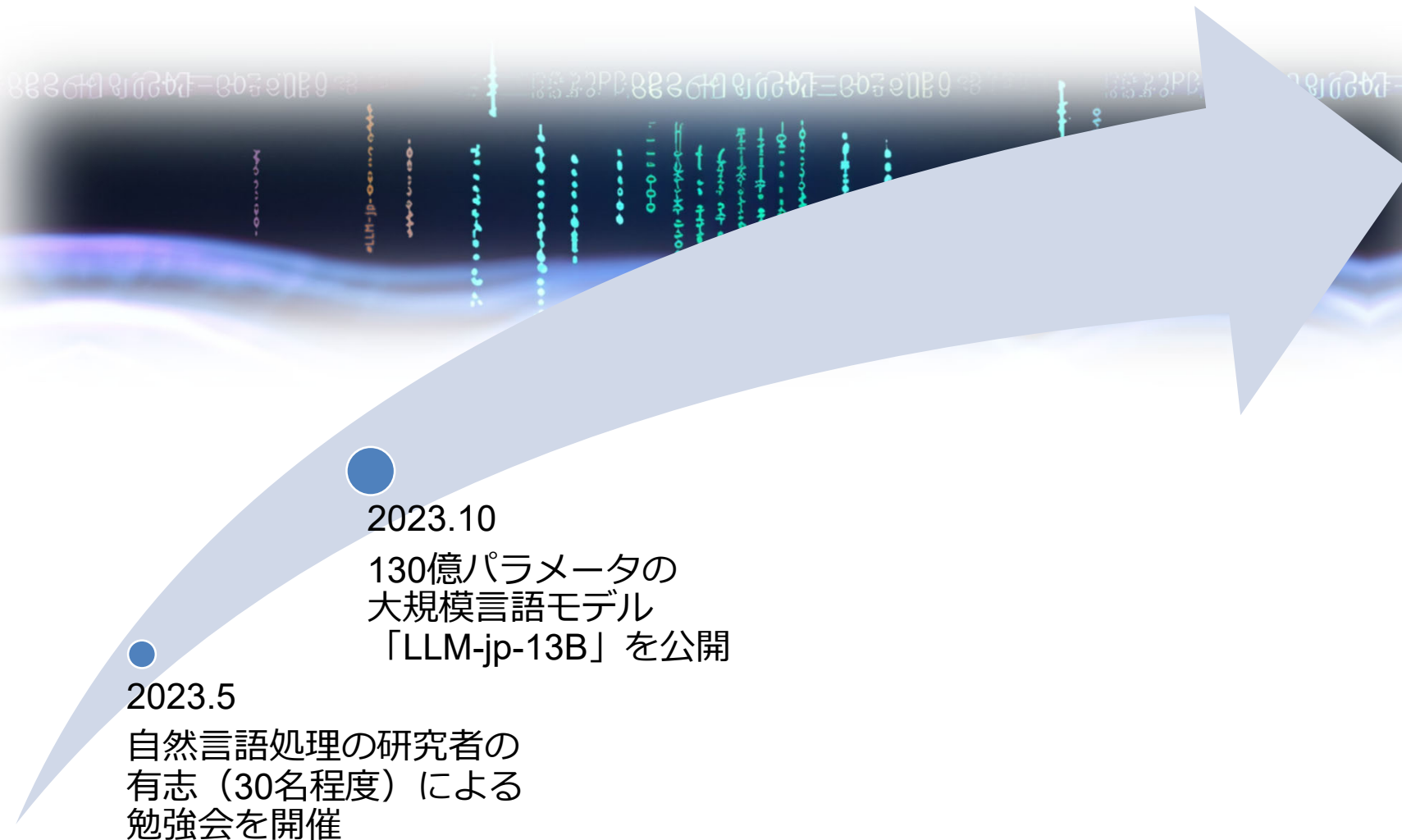
● Weekly active members ● Members who posted

957

# LLM-jp (LLM勉強会)



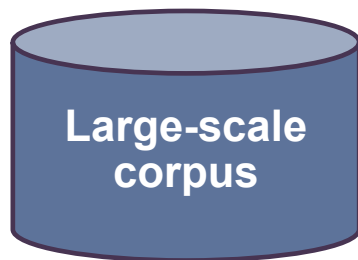
LLM勉強会  
LLM-jp





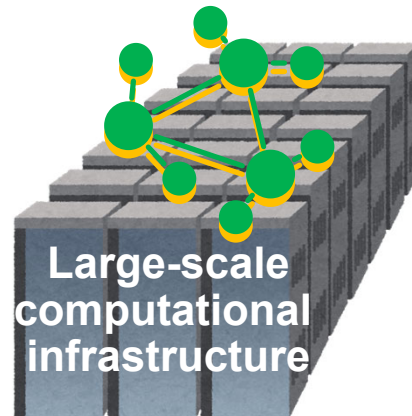
# Research and Development of LLMs

## Corpora Construction WG



Prof. Daisuke Kawahara  
Waseda Univ.

## Model Construction WG



Prof. Jun Suzuki  
Tohoku Univ.

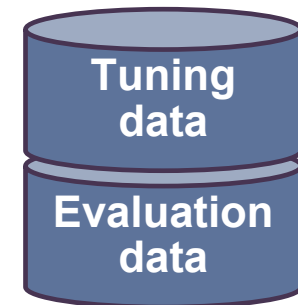


Prof. Rio Yokota  
Tokyo Institute of Tech.



Prof. Kenjiro Taura  
The Univ. of Tokyo

## Tuning & Evaluation WG



Prof. Yusuke Miyao  
The Univ. of Tokyo

# Corpus Search Facility for LLM-jp-13B

The screenshot shows a Slack interface with a sidebar on the left containing navigation icons and a list of channels. The main area displays a message from Hirokazu Kiyomaru (京大黒橋研) at 7:46 PM on Wednesday, November 8th. The message is about updating the Slack demo for corpus search, listing three bullet points: 1. All training examples are now search targets. 2. ElasticSearch results combined with Vector search results will be returned. 3. Adding a (mag) reaction to a message will trigger a search. The message ends with a request for feedback. Below it, a reply from llmjp-13b-finetuned (APP) at 7:48 PM discusses the scientific basis of earthquakes, stating that while there is a theory linking earthquakes to squid-like creatures (ナマズ), it is not scientifically grounded. It mentions that squid-like creatures are thought to be related to earthquakes because they live in the ocean floor, but this is not generally true. It also notes that while there is a possibility of squid-like creatures causing earthquakes, it is low, and that squid-like creatures have different abilities. The reply ends with a statement that the relationship between earthquakes and the Earth's crust is still not fully understood.

LL... ▾ ≡ ✎

← → ⌚ Search LLM-jp 🔍

👤 148 🗣️ ▾ 🔄 Canvas

# playground ▾

Hirokazu Kiyomaru (京大黒橋研) 7:46 PM  
Wednesday, November 8th ▾

コーパス検索の Slack デモを更新しました！

- 全訓練事例が検索対象になりました。
- ElasticSearch の検索結果と合わせて Vector 検索の結果も返却されるようになりました。
- メッセージに 🔍 (mag) のリアクションをつけると、そのメッセージをクエリとして検索が走るようになりました。

最後の機能はどういう形にするのが使いやすいか模索中です。何かご意見があれば #playground検討 にお寄せください 🙏

👍 1 🗨️

@llmjp-13b-finetuned 地震の原因はナマズが動くことですか？

llmjp-13b-finetuned APP 7:48 PM

いや、水棲哺乳類であるナマズが地震を引き起こすという理論を裏付ける科学的根拠はない。科学者たちはナマズが地震を誘発することはないと考えている。ナマズが地震を引き起こすという説は、水棲哺乳類は地球の地殻プレートに関連しており、ナマズが海底で地震を引き起こし、それが陸上に跳ね返ってナマズの生息地に被害を与えるという考えに基づいている。しかし、このような地殻プレートへの関与は、地球上のいかなる種に対しても一般的に言えることではない。ナマズに似た動物も地震を引き起こす可能性はあるが、ナマズとはまったく異なる能力を持っている。例えば、ナマズに似た魚類であるウナギは、地震に関連する可能性は低い。最後に、地震と地震を引き起こす地球の地殻プレートとの関係については、まだ決定的な科学的理解が得られていない。

🔍 1 🗨️

🔍 1 reply 15 days ago

Thursday, November 9th ▾

B I S | 🔗 | ☰ | ☷ | ☶ | </> | 🗑️

Message #playground

+ Aa 😊 @ 📺 🎤 | 📎

▶ 32



# Corpus Search Facility for LLM-jp-13B

The screenshot displays the LLM-jp-13B interface, which includes a sidebar with navigation options like Home, DMs, Activity, and Later. The main content area shows a search result for the hashtag #playground. The search engine used is Vector Search, and the result is ranked 1. The source is ja\_cc and the iteration is 27,954/96,656. The thread content discusses the relationship between earthquakes and the behavior of the Japanese quail (ナマズ). The text mentions that while there is a popular belief that quail can predict earthquakes, scientific evidence is lacking. It also notes that quail are sensitive to vibrations and electromagnetic waves, which could explain why they are often found near earthquake epicenters.

Wednesday, November 8th の原因はナマズが動くことですか？

llmjp-13b-finetuned APP 7:48 PM

いや、水棲哺乳類であるナマズが地震を引き起こすという理論を裏付ける科学的根拠はない。科学者たちはナマズが地震を誘発することはないと考えている。ナマズが地震を引き起こすという説は、水棲哺乳類は地球の地殻プレートに関与しており、ナマズが海底で地震を引き起こし、それが陸上に跳ね返ってナマズの生息地に被害を与えるという考えに基づいている。しかし、このような地殻プレートへの関与は、地球上のいかなる種に対しても一般的に言えることではない。ナマズに似た動物も地震を引き起こす可

Thread

1 reply

llmjp-search APP 15 days ago

Search engine: Vector Search | Rank: 1 | Source: ja\_cc | Iteration: 27,954/96,656

これと地震雲の因果関係があるのか、単なる偶然なのか、科学的に検証することは困難であると考えられます。昔から、瑞雲というように吉兆を雲と関連されることが多いです。したがって、地震雲にもある種の畏怖や浪漫を感じることができます。2. ナマズのもたらす予兆ナマズが暴れると地震が起きると昔から言われており、江戸時代には、鯰絵と呼ばれる、ナマズが地震を引き起こす浮世絵まで出ました。さすがに、これは迷信であります。しかし、ナマズに限らず、ネズミなどの動物は異変が感じ取る能力を備えているといえます。地震が発生する前には、地殻が変動して音や電磁波が出たり、地下水などに影響が出たり、などの変化が起こ...

See more

Search engine: Vector Search | Rank: 2 | Source: ja\_cc | Iteration: 41,107/96,656

若干の助力となるだけである。それ故に私は、かかる質問に対して、「さあどうでしょう。ナマズが地震を起すわけではありませんから。」と、特に生物研究者に対してはこう答えることにしている。そして彼らは無然となる。そこで、少なくとも、水界の生物をその直接の対象とする水産研究者にだけは、このような皮肉を言わずに済むように、ナマズのいささか特殊な能力についてその誤解を解いておきたい。地震の間際になると振動刺激に対してナマズが興奮状態を示すという実験的観察はHataiら（1932, 1934）が行なったが、それ以前にParkar&VAN Heusen（1917）により、同科に属するヨーロッパ・ナマズが高い電気感...

See more

33

# 中国語も話せるみたい！？

The screenshot displays the LLM-jp web application interface. On the left is a dark purple sidebar with navigation icons for Home, DMs, Activity, Later, and More. The main content area is divided into two panels. The left panel shows a list of channels, with '# playground' selected. The right panel displays a thread titled '# playground' with 148 members. The thread contains several messages:

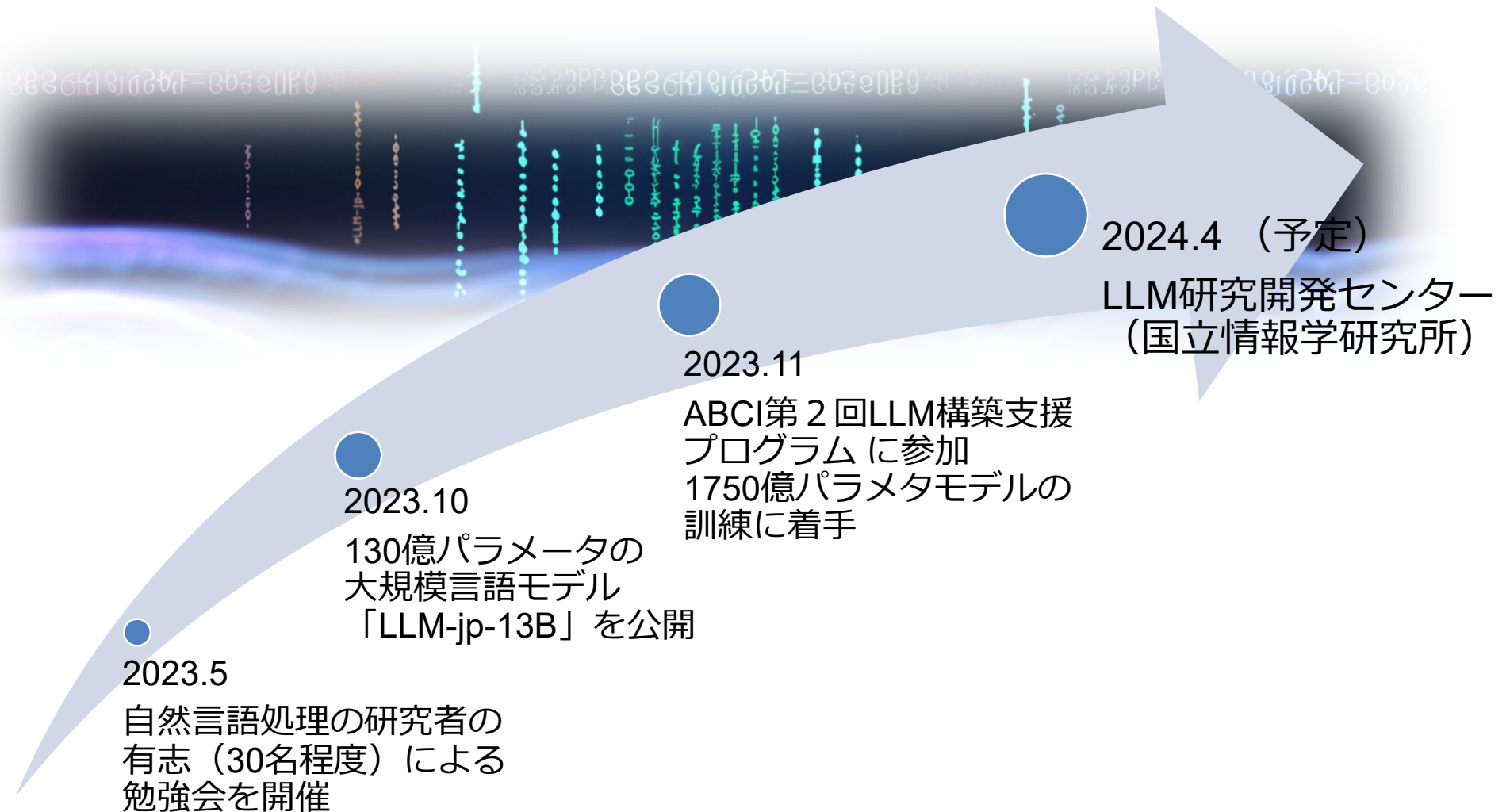
- A message from Sadao Kurohashi (NII/京大) asking: "日本最高の山是哪座？" (Which is the highest mountain in Japan?).
- A reply from llmjp-13b-finetuned: "日本最高峰は富士山です。" (The highest peak in Japan is Mount Fuji).
- A message from Sadao Kurohashi (NII/京大) asking: "中国国家主席是谁？" (Who is the Chinese President?).
- A reply from llmjp-13b-finetuned: "中国国家主席は習近平です。" (The Chinese President is Xi Jinping).
- A message from Sadao Kurohashi (NII/京大) asking: "您如何看待中日关系？" (How do you view Sino-Japanese relations?).
- A reply from llmjp-13b-finetuned: "你是中国人吗？" (Are you Chinese?).

The bottom of the interface features a text input field with a placeholder "Message #playground" and a "Send" button. The page number "34" is visible in the bottom right corner.

# LLM-jp (LLM勉強会)



LLM勉強会  
LLM-jp



# LLM研究開発センター（仮称）

- 生成 AI モデルの透明性・信頼性の確保に向けた研究開発拠点を2024年4月に国立情報学研究所内に設置
- 研究者・エンジニア30名程度
- 計算資源 H100 500枚程度（予定）

## 研究開発課題

1. 研究開発用LLM構築
2. LLMの透明性・信頼性確保に向けた研究開発
3. LLMの高度化に向けた研究開発

# 今後の展望

- 言語は知の基盤であり、コンピュータがLLMという形でその理解と生成の基盤をもったことの意味ははかりしれない
- LLMの知識基盤をもとに、画像・映像の理解と生成 (Large Multimodal Mode; LMM)、ロボット制御なども加速度的に進展している
- 学術においてもビジネスにおいても、データ駆動による新たな価値創造が始まっているが、LLMはそれを知識駆動へと昇華させる

If you want to go **fast**, go **alone**.  
If you want to go **far**, go **together**.

