

大規模言語モデル研究開発センター (LLM研究開発センター: LLMC)



どんな研究？

ChatGPTなどの大規模言語モデル(LLM)が社会基盤として浸透するうえで多くの課題があります。オープンかつ日本語に強いLLMを開発し、その原理解明やモデルの公開を通して、LLMの透明性や安全性の確保などの最重要課題に取り組みます。

何がわかる？

LLMの事前学習用コーパスと性能との関係、性能評価の方法(日本語の理解・生成能力を含む)、チューニングによる改善などの多くの知見により、LLMCが公開したモデル／データを用いた検証や、特定分野向けのモデル構築が可能になります。

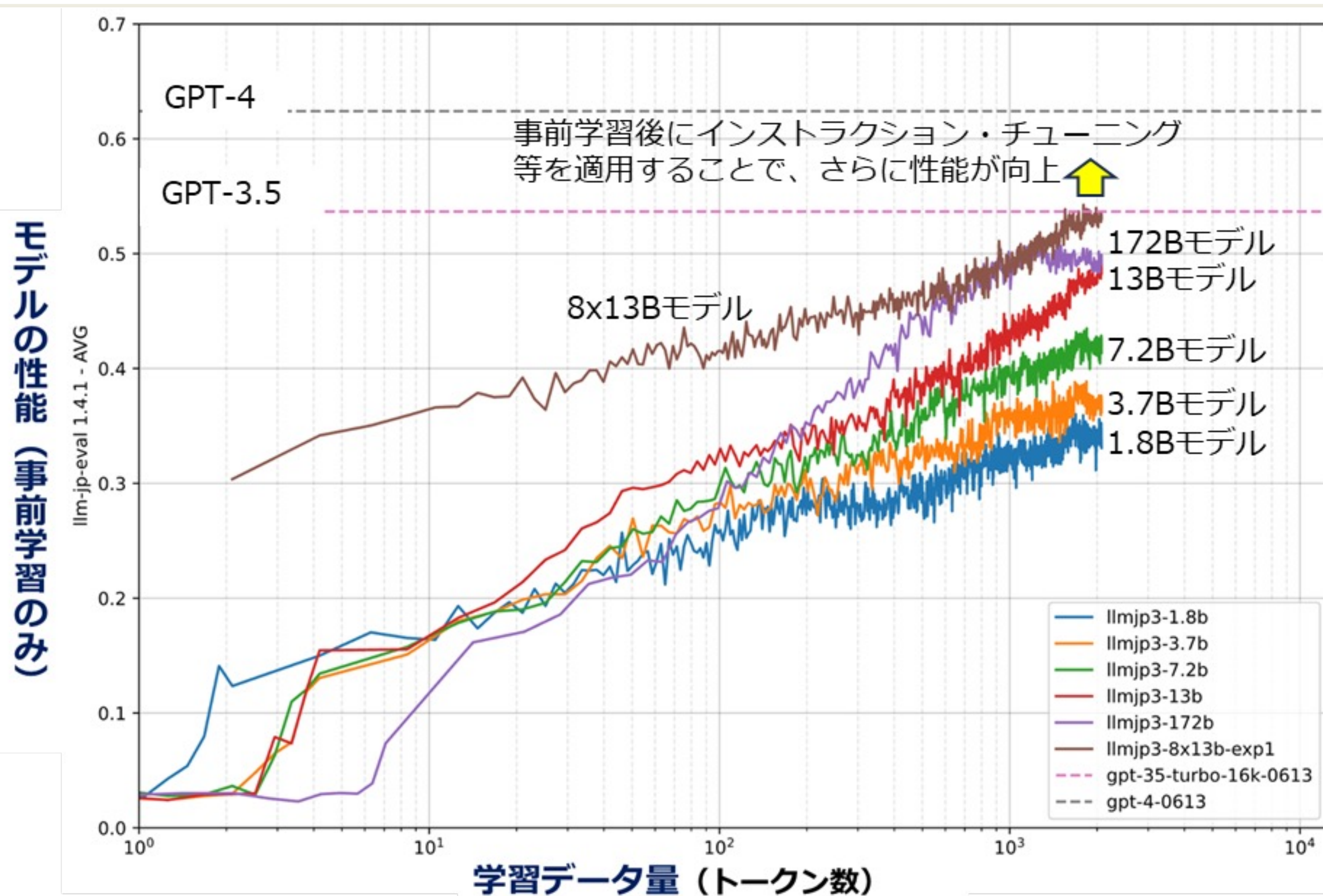
研究内容

LLM研究開発センターでは以下の研究開発を行っています。

- 研究開発用LLM構築**
コーパス整備、計算環境整備、評価用ベンチマーク作成などを行うとともに研究開発用のLLMを構築します
- 透明性・信頼性確保に向けた研究開発**
生成AIの挙動原理を解明すること、またデータ改変やデータバイアス等の影響を抑制する技術を開発することなどにより、生成AIの透明性・信頼性を確保します
- LLMの高度化に向けた研究開発**
ドメイン適応、モデル自体の軽量化など、生成AIモデルの高度化に資する研究開発を行います

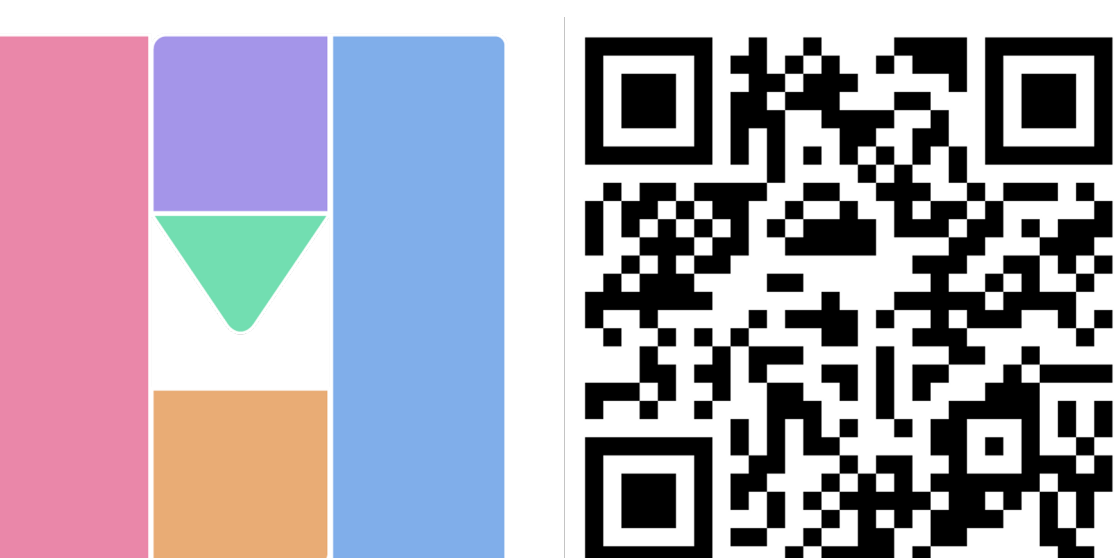


これまでに公開したモデルの性能



Q: この画像には何が写っていますか？
A: 画像には、白と赤のバスが都市の通りを走っている様子が写っています。バスは比較的大きく、広告が後部に貼られているようです。また、周囲には多くの人や建物があります。

マルチモーダルデータを扱うモデルLLM-jp-3-VILA-14Bは、日本語タスク評価で一部OpenAI GPT-4o に匹敵する性能を達成



モデル、データ、ツールをLLM-jpのサイトで公開中

LLM-jp: 自然言語処理及び計算機システムの研究者を中心として、大学・企業等から2,200名以上(*)が集まり、ハイブリッド会議、オンライン会議、Slack等を活用してLLMの研究開発について情報共有を行うとともに、共同でLLM構築等の研究開発を行っています

(*)2025年4月1日時点



LLM研究開発センター(LLMC)の研究紹介

小田 悠介、劉 超然、小林 和馬、清丸 寛一、磯沼 大、児玉 貴志、中山 功太、劉 倩瑩、
Nguyen Ha Thanh、Yang Zhishen、Su Myat Noe、橘 秀幸、江 俊鋒 (LLMC 特任研究員)



オープンな日本語大規模言語モデルの開発取組

LLMCにおけるオープンな大規模言語モデルの開発取組について紹介する

小田 悠介

推論タスクに強い大規模言語モデルの開発

推論タスクに強い大規模言語モデルの開発取組について紹介する



劉 超然



医療大規模言語モデル

医療従事者が安心して利用できる国産LLM基盤について紹介する

小林 和馬

大規模言語モデルの学習コーパスの分析

大規模言語モデルの振る舞いを学習コーパスの側面から分析する
取り組みを紹介する



清丸 寛一



逆学習による大規模言語モデルの安全性向上・理解

大規模言語モデルの逆学習により、有害な生成を抑制する研究や、
重要な学習データを特定する取り組みを紹介する

磯沼 大

日本語大規模言語モデルの有用性と安全性の両立

日本語 LLM の安全性向上を目的とし、有用性を損なわずにチューニングする手法について紹介する



児玉 貴志



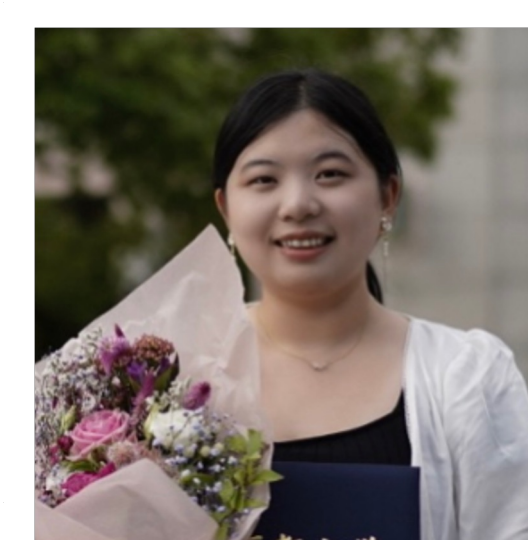
llm-jp-judge: 日本語LLM-as-a-Judge評価ツール

LLMの多面的な分析を容易にするため LLM-as-a-Judge 評価を
統一的に扱うことができるツールを提案する

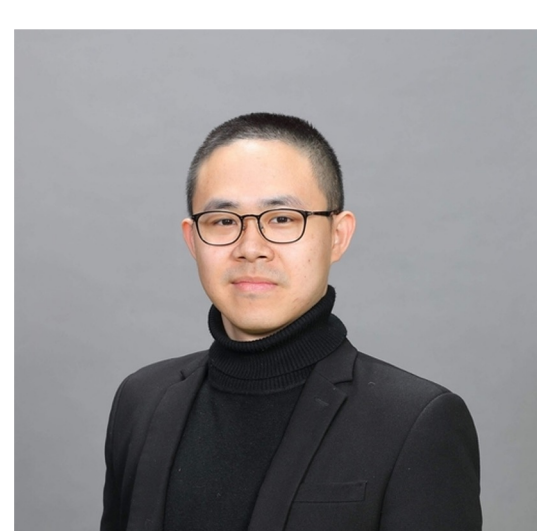
中山 功太

What language do LLMs think in?

I investigate whether LLMs trained with substantial
Japanese data exhibit higher probabilities for Japanese in
their intermediate layers



劉 倩瑩



LLMs for Legal Reasoning

I will explain my research on leveraging LLMs for legal
reasoning, extending to broader high-stakes domains

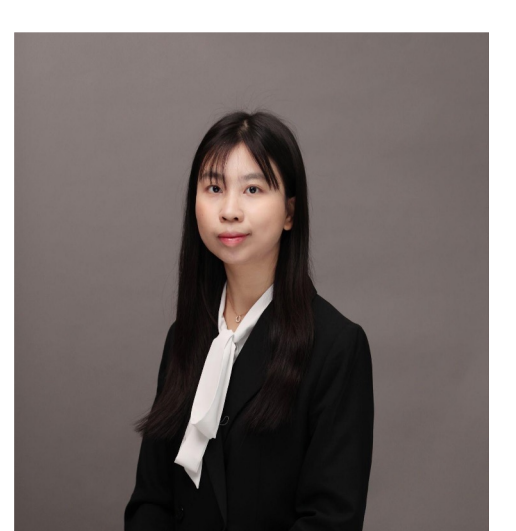
Nguyen Ha
Thanh

LLMs Beyond Text: Multimodal Understanding

Advancing multimodal large language models that integrate
text, image, video, and domain-specific knowledge to
enhance understanding and generation across diverse tasks
and domains



Yang Zhishen



Multimodal Models and Computer Vision

Research focuses on multimodal AI systems, particularly in
the areas of vision-language modeling, Video Question
Answering (VideoQA), and real-world video understanding

Su Myat Noe

音声技術と大規模言語モデル

大規模言語モデルによる音声言語処理について紹介する



橘 秀幸



Scientific Documents and LLMs

I will explain my research on LLMs for scientific documents
and discovery

江 俊鋒