



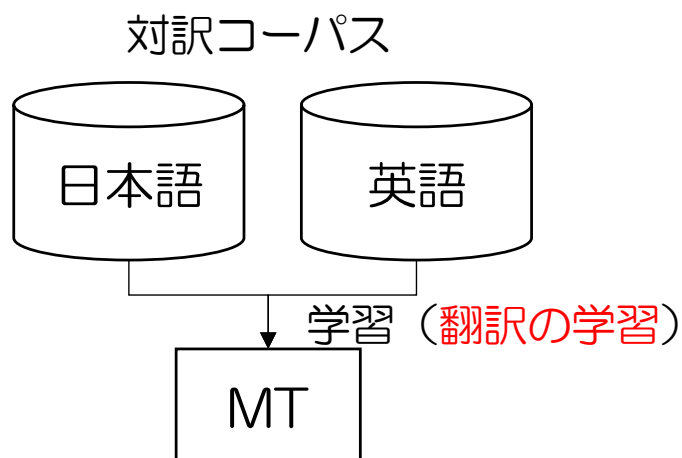
# MT Summit 2023 参加報告

AAMT理事 / NICT 田中英輝

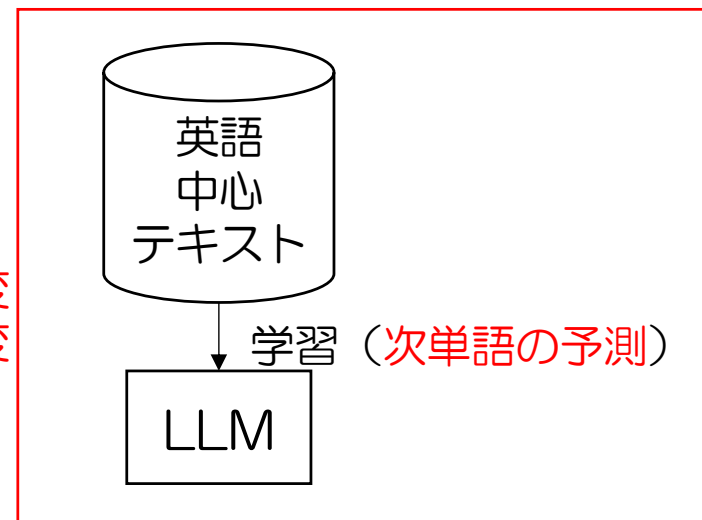
# 会議の概要

場所	中国、マカオ Studio City ホテル
期間	9月4日 - 5日 ワークショップ4件 9月6日 - 8日 本会議
形式	ハイブリッド
参加人数	169名、内、現地参加110名程度
参加国	25カ国
スポンサー	16社
招待関連	基調講演1、招待講演2件、招待セッション2
研究発表	リサーチトラック 33件 ユーザーズトラック 21件

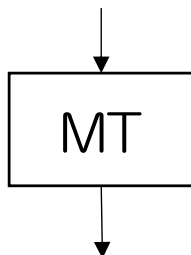
# 機械翻訳(MT)と大規模言語モデル(LLM)



巨大  
学習が大変  
変更も大変



I went to the bank yesterday.

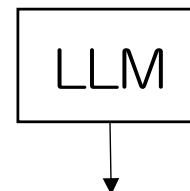


昨日、私は銀行に行きました。

入力に対する「翻訳」

次の英語を日本語に翻訳してください

I went to the bank yesterday. ○○



昨日、私は銀行に行きました。

入力に対する「次単語」の予測  
翻訳に限らない汎用システム

# 招待講演者のLLM翻訳へのコメント



基調講演

Prof. Min Zhang



ほぼ英語でしか学習していないの翻訳可能！  
多言語翻訳に対応  
対話的に翻訳可能  
文書単位での翻訳可能  
受け付ける文のスタイルが多様



従来手法での自動評価が困難  
専門用語が苦手  
計算コストが大きい  
低資源言語の翻訳は苦手



招待講演

Dr. Ondrej Bojar



学習データが不明のため正確な評価ができない  
LLMにもできないことは多い  
研究の多様性確保が重要. LLM一極集中はダメ  
Help ChatGPT fall slowly and gracefully



招待講演

Dr. Dorothy Kenny



目的に合わせた翻訳が可能

# 従来システムとの比較 「単純な翻訳」の場合

Comparing Chinese - English MT  
Performance Involving ChatGPT and MT  
Providers and the Efficacy of AI mediated  
Post - Edition, Larry P Cady +

# 研究概要

- 機械翻訳システムの中英、英中翻訳の比較評価
  - システム：
    - Baidu
    - DeepL
    - Google Translate
    - NiuTrans
    - Youdao
    - ChatGPT3
    - ChatGPT4
  - 評価：特許（バイオ、計算機工学、電子工学）  
各3,000文の翻訳をBleuで評価

# 同一文に対するGoogleとGPT-4のBleuスコア散布図例

中→英翻訳

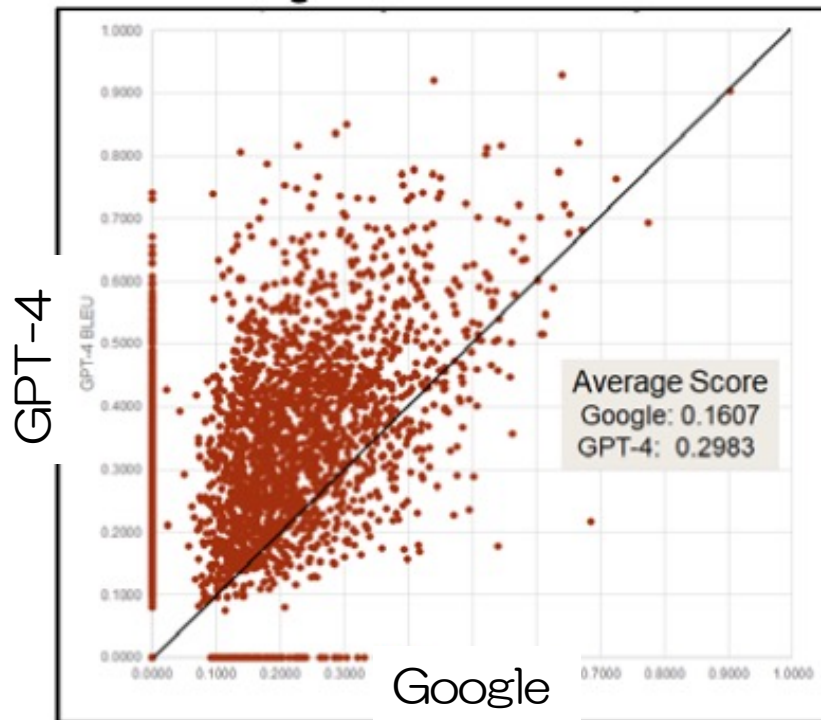


Figure 2A. Google vs GPT-4 :  
Chinese to English

英→中翻訳

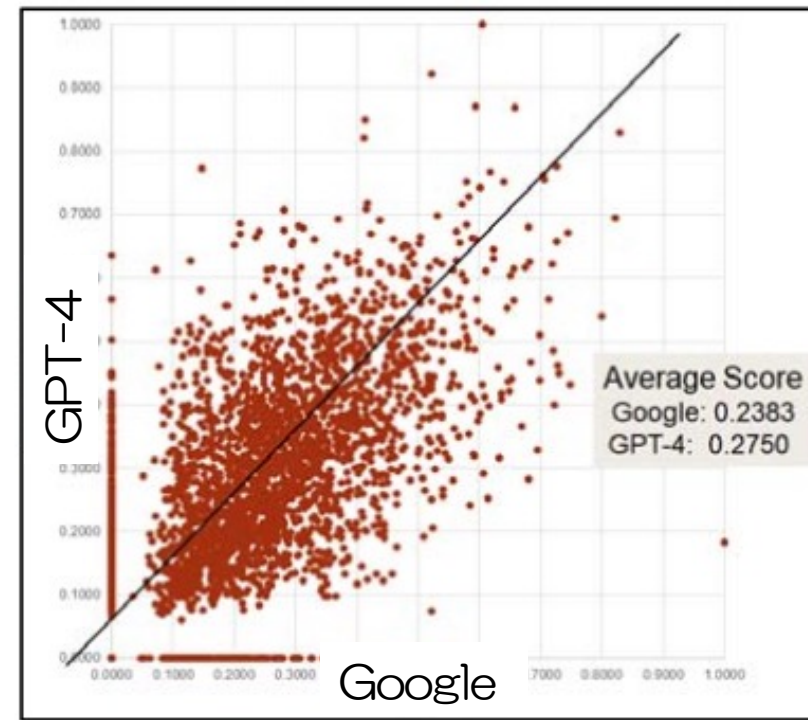


Figure 2B. Google vs GPT-4 :  
English to Chinese

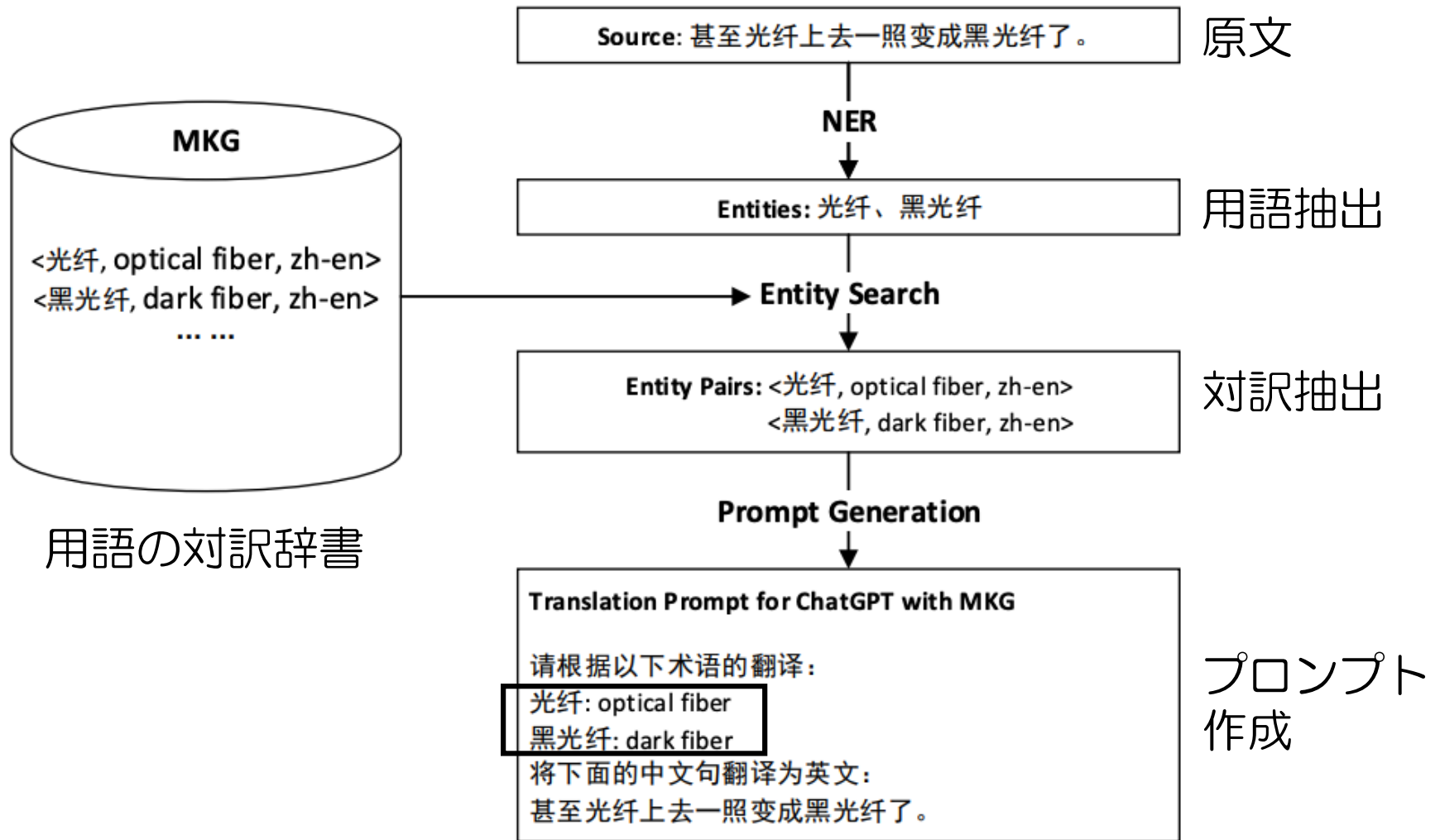
すべてにおいて圧倒的なシステムはない

# プロンプトの利用 専門用語翻訳

Leveraging Multilingual Knowledge Graph  
to Boost Domain-specific Entity  
Translation of ChatGPT, Min Zhang+



# MKG (対訳辞書) を利用した翻訳



# 実験

## システム

---

比較	Google Translate
	ChatGPT (gpt-3.5-turbo)
提案	ChatGPT (gpt-3.5-turbo) + 辞書 (MKG)

---

System	バイオ			光通信			無線通信		
	BLEU	ACC	COMET	BLEU	ACC	COMET	BLEU	ACC	COMET
ChatGPT	29.37	0.500	81.2	22.10	0.563	81.6	21.08	0.507	80.5
提案 ChatGPT+MKG	35.58	<b>0.932</b>	<b>82.8</b>	25.23	<b>0.865</b>	<b>82.6</b>	<b>32.33</b>	<b>0.886</b>	<b>82.6</b>
Google Translate	<b>37.06</b>	0.527	81.6	<b>29.80</b>	0.619	81.6	26.13	0.529	80.0

ACC 用語正解率

- 提案手法とChatGPTの比較
  - すべての指標で提案手法が優位
- 提案手法の用語正解率（ACC）の高さが顕著

# 無線通信分野の実例

Type	Content
Source	电源线与空开连接器接触不良，如电源线剥线过短导致压绝缘皮等。
Reference	The connection between the <b>power cable</b> and <b>circuit breaker connector</b> is poor. For example, the <b>insulation sheath</b> is pressed because the stripped length is too short for the <b>power cable</b> .
Entities from MKG	<空开连接器, circuit breaker connector, zh-en> <绝缘皮, insulation sheath, zh-en> <电源线, power cable, zh-en>
ChatGPT	Poor contact between <b>power cord</b> and <b>circuit breaker connector</b> , such as insufficiently stripped wires of the <b>power cord</b> causing pressure on the <b>insulation</b> .
ChatGPT+ MKG	The <b>power cable</b> has poor contact with the <b>circuit breaker connector</b> , which may be caused by the stripping of the <b>power cable</b> being too short and pressing against the <b>insulation sheath</b> .
Google Translate	The <b>power cord</b> is in poor contact with the <b>circuit breaker connector</b> , for example, the stripping of the <b>power cord</b> is too short and the <b>insulation</b> is pressed.

Table 6: A case from the wireless domain, which contains four domain entities.

赤：誤り 青：正解

# プロンプトの利用 修正命令による個別ドメイン対応

Human-in-the-loop Machine Translation  
with Large Language Model, Xinyi Yang+

# 概要

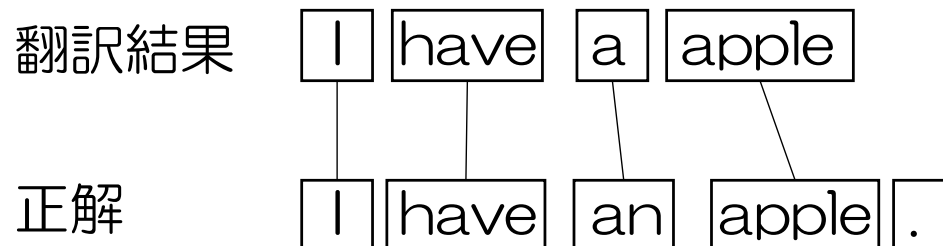
## □ 動機

- LLMでの専門分野向けのファインチューニング（学習）は困難

## □ 提案手法

- LLMによる翻訳
- 誤りの人手修正 → 修正命令（プロンプト）の自動作成
- 修正命令の蓄積によるデータベースを構築
- 修正命令の検索と人手修正によって専門分野翻訳を実現

## 修正命令の自動作成

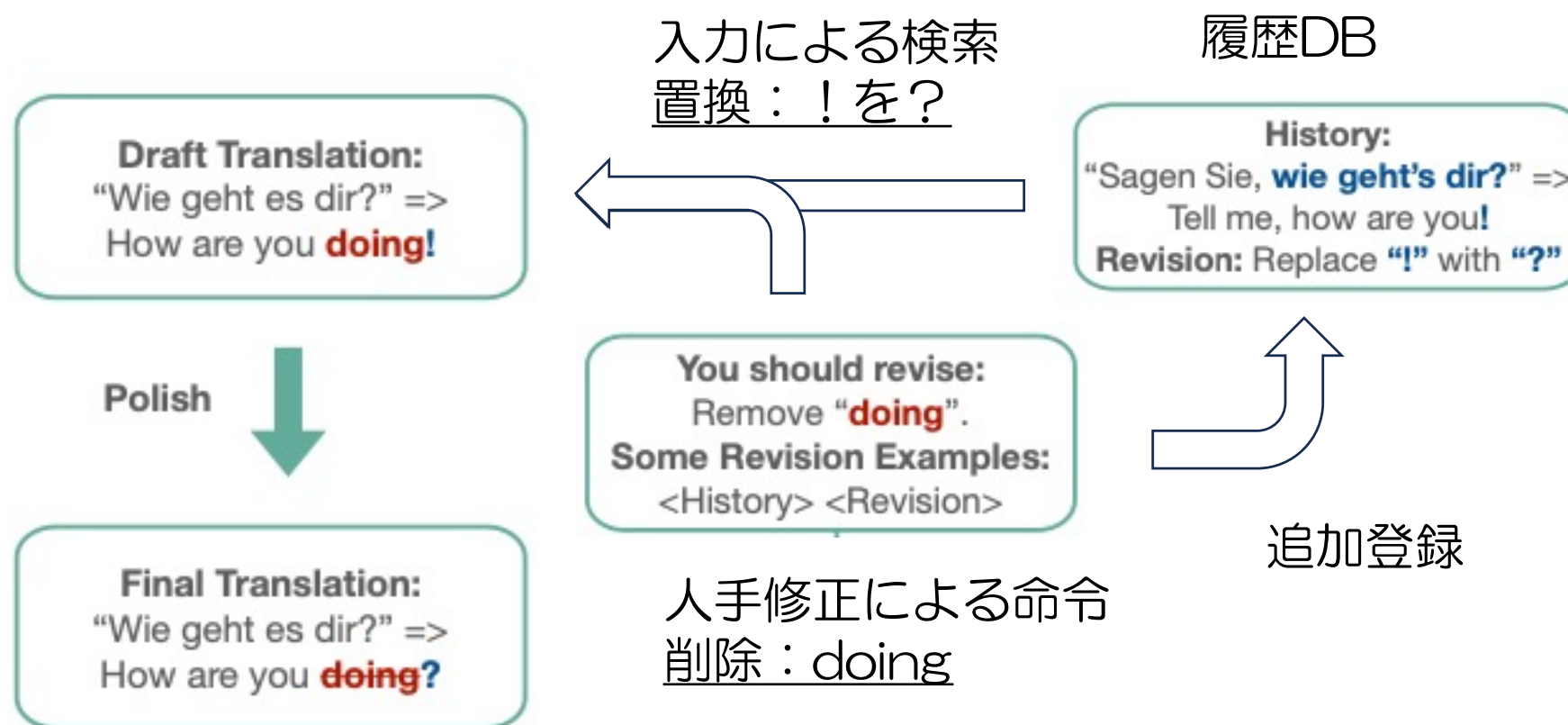


修正命令

置換：aとan

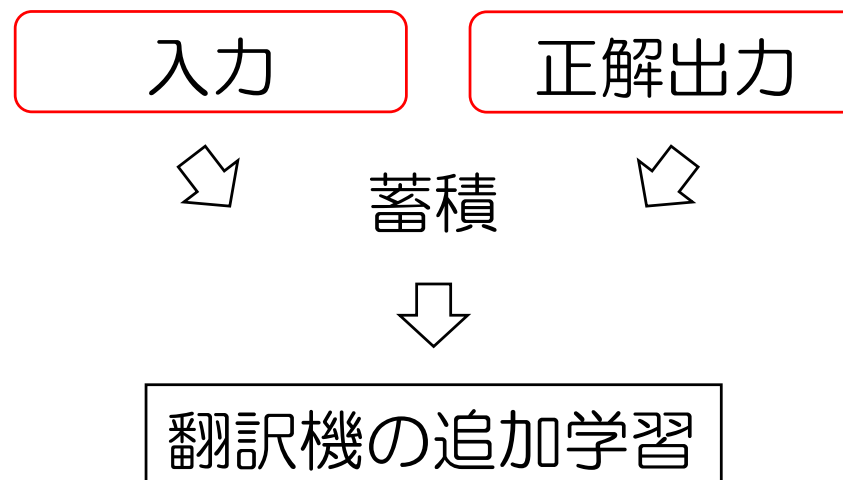
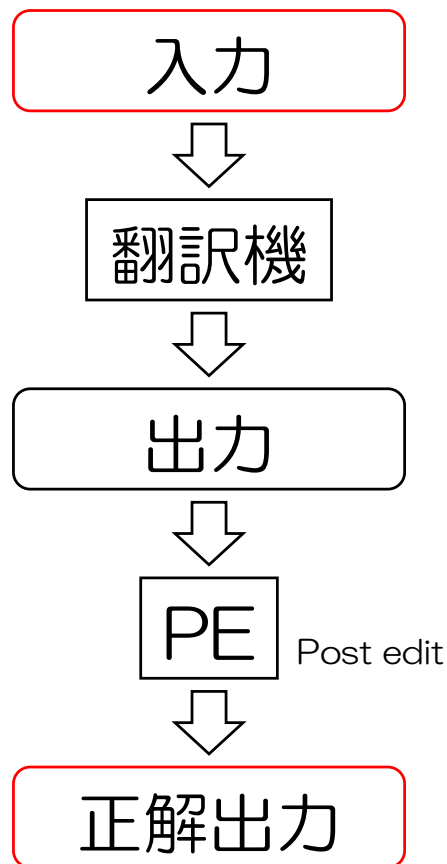
挿入：appleの後に“.”

# 動作例（独→英翻訳）



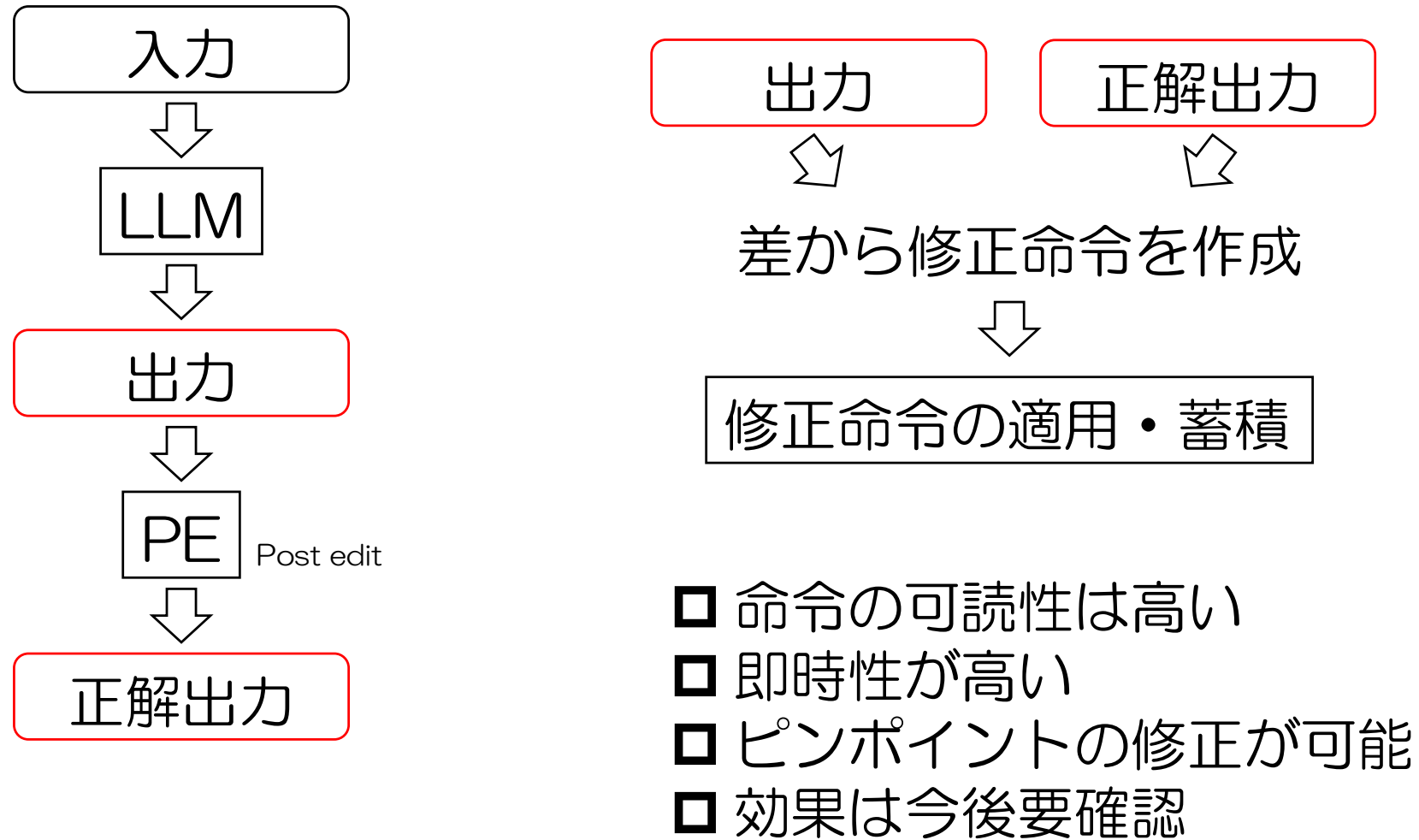
5分野での独英翻訳シミュレーション実験  
顕著な翻訳性能向上（Bleu）は見られなかった

# 従来の専門分野適応化



- 性能向上は顕著
- 再学習までの時間がかかる
- ピンポイントの修正は難しい

# 本提案の専門分野適応化





# プロンプトの利用 翻訳目的の指定

Optimizing Machine Translation through  
Prompt Engineering: An Investigation into  
ChatGPT's Customizability, Yamada

# 概要

## □ 動機

- 目的に合わせた翻訳を実現したい
  - 翻訳には目的・仕様がある (ISO17100)

## □ 研究の目的

- プロンプトにより目的を指定した翻訳の調査
- 使用システム

---

	DeepL	
比較	Google Translate	
	ChatGPT Plus (GPT4)	(翻訳命令)
	ChatGPT Plus (GPT4)	(翻訳命令+目的) 1
提案	ChatGPT Plus (GPT4)	(翻訳命令+目的) 2
	ChatGPT Plus (GPT4)	(翻訳命令+目的) 3

---

- 目的：マーケティング、文化背景、動的等価性



# 翻訳結果：文化的背景を考慮した慣用句

[source text] 私たちは同じ釜の飯を食べた仲です。

Type	Target sentence	C.S.	Rank
DL	We are friends who <u>ate out of the same pot.</u>	0.772	1
GT	We <u>ate rice from the same pot.</u>	0.727	5
GPT	We <u>ate rice from the same pot.</u>	0.727	5
v1	We have <u>shared the same pot of rice.</u>	0.743	4
v2	We have <u>been through thick and thin together.</u>	0.759	2
v3	We've <u>broken bread together.</u>	0.744	3

目的プロンプトあり

v2の翻訳が最良

# 目的：動的等価性\*の保存

\*Dynamic equivalence

## □ 目的に合わせたプロンプト（下線は英語で記載）

動的等価性とは、…（省略）…原文の「子羊」と訳文の「アザラシ」は、読者に同じ反応を引き起こすと考えられます。

[source text] Lamb of God（神の子羊）  
[target text] Seal of God（神のアザラシ）

このコンセプトと例にしたがって、次の〔原文〕を動的等価表現を使って英訳してください。英語圏の文化圏で理解されるような訳に置き換えてください。

[source text] 彼女の歌声は美空ひばりを彷彿とさせる。

# 翻訳結果：動的等価性の保存

[source text] 彼女の歌声は美空ひばりを彷彿とさせる。

Type	Target sentence	C.S.	Rank
DL	Her singing voice is reminiscent of Hibari Misora.	0.876	1
GT	Her singing voice is reminiscent of Hibari Misora.	0.876	1
GPT	Her singing voice reminds me of Misora Hibari.	0.873	2
v1	Her singing voice evokes memories of Judy Garland.	0.830	3
v2	Her singing voice is reminiscent of Billie Holiday.	0.823	5
v3	Listening to her sing, one can't help but think of Ella Fitzgerald.	0.826	4

目的プロンプトあり

v1-v3は原文と違う歌手が出現  
最適な訳は翻訳者が検討すべき

# トピックス

- IAMT Award of Honor
  - Dr. Qun Liu, Huawei Noah's Ark Lab.
- Mtsummit 2025
  - スイス、ジュネーブで開催



# まとめ

- 会議の概要
- 大規模言語モデルに関する話題
  - 招待講演者
  - 従来システムとの比較 単純なLLM翻訳
    - 大差なし
  - プロンプトの利用
    - 専門用語翻訳
    - 修正命令による個別ドメイン対応
    - 翻訳目的の指定
- トピックス
  - 次回開催 2025年ジュネーブ
- 予稿集はACL anthology
  - <https://aclanthology.org/events/mtsummit-2023/>