

# MANTRA

Delivering Manga across language barriers.

Mantraのミッション

世界の言葉で、マンガを届ける。

## 創業

2020年1月28日

## 事業内容

マンガに特化した  
機械翻訳技術の研究開発

## 主な取引先 (敬称略)

株式会社集英社,  
株式会社小学館,  
大日本印刷株式会社,  
NAVER WEBTOON Ltd.  
他多数



## 出資者

集英社 小学館

 KADOKAWA

SQUARE ENIX

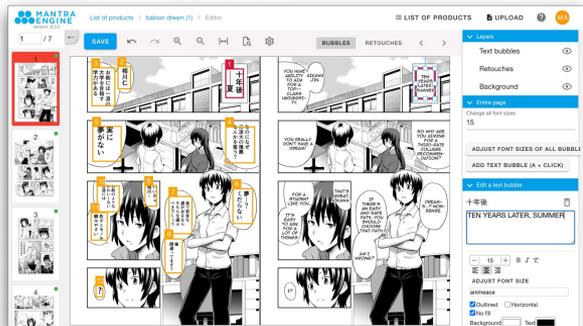
 東大IPC

Innovation Platform for The University of Tokyo

DEEP(ORE)

 m)POWER など

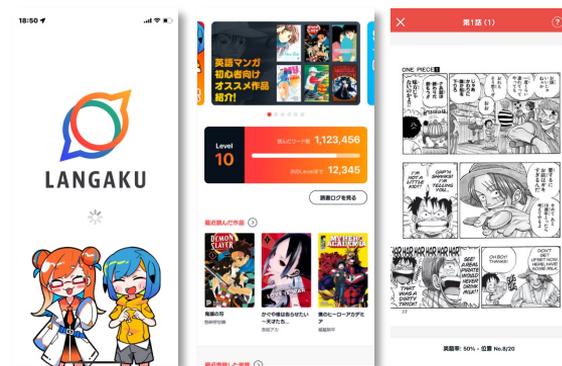
## マンガ用AI翻訳ツール Mantra Engine



## マンガ・コミック・ 小説の翻訳サービス



## マンガで英語学習サービス Langaku (ランガク)



# ● 背景：世界で急成長する“Manga”

Confidential



北米：4年で売上5x<sup>[1]</sup>

- “Manga”は大人向け書籍の売上ランキング3位

フランス：

4年で売上2x以上<sup>[2]</sup>

[1] <https://www.npd.com/wp-content/uploads/2023/03/US-Books-MangaCampaign-infographic.pdf>

[2] <https://www.gfk.com/fr/insights/dynamique-roman-premier-trimestre-2023>

画像1：<https://twitter.com/CoveredGeekly/status/1712502282640580659/photo/1>

画像2：<https://nerdreactor.com/anime-expo-2023-cosplay-wrestling-federation-late-comedy-showdown/>

画像3：<https://twitter.com/TamashiiNations/status/1675173514226307073/photo/1>

# ● マンガ海外展開の現状 | 概観

翻訳出版されるマンガは全体のごく一部

- アメリカ約9%<sup>\*1</sup>、フランス約10%<sup>\*1</sup>、台湾18%<sup>\*2</sup>

多くの作品が日本の数ヶ月~数年遅れで出版

- 多言語同時配信される作品は極めて限定的
- 同時配信対応される言語も限定的

→ 依然、マンガの海外海賊版ユーザが後を絶たない

# ● マンガの海外展開の現状 | 翻訳プロセス

Confidential



- 基本的に全て人手作業
- 煩雑なデータやり取り



- 多くの作品は翻訳されず
- される場合でも時間がかかる



保存

画像を開く

設定



翻訳・文字組み・校正の作業を効率化し、マンガの海外版制作までの時間をおよそ半分までにまで圧縮します。

## ページ全体

すべてのフォントサイズを変更

15

フォントサイズの自動調整

吹き出しの追加

## 吹き出しの編集

これは夢への第一歩

THIS IS THE FIRST STEP TO MY DREAM.

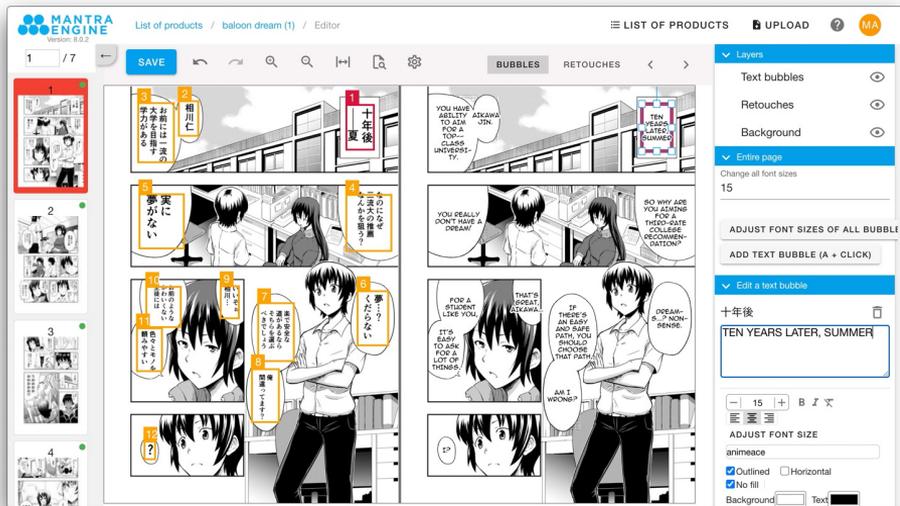
- 16 + B I X 自動調整 animeace

 縁取り  横書き 塗り潰さない | 背景色            文字           

Toよりも Towardsの方が適切かと思ます。(校閲者)

# Mantra Engine: マンガAI翻訳ツール

Confidential



対応言語数  
19言語

翻訳ページ数  
月次10万ページ



# Mantraのマンガ処理AI技術

## テキスト抽出

5. うおおおおお おおおお!?

1. ...天文部

2. ここか

3. 失礼しまーす

4. ん?

ガラッ

## 翻訳

あ

A

中

B

## 画像生成

レタッチ

組版

sniff sniff waaaah

WHIPPED UP WITH GOLD AND NOW IT'S ALREADY TIME TO LAUNCH AGAIN.

IN THE END I COULDN'T FIND THE PROTOTYPE...

# ● マンガ翻訳の難しさ

## 「面白い」翻訳

意味が伝わるだけではダメ  
作品として読んで面白い翻訳が必要

## 評価の難しさ

自動評価では人間と同じ精度  
プロ翻訳者の主観的な評価に依存

## 多様な文脈

前後の文、ストーリー、キャラ、  
文化的背景など多様な文脈が存在

## マルチモーダル

画像情報で訳し方が変わる  
キャラクターの口調や感情など



# 「良い翻訳」とは…?

Confidential

原言語:

「シアワセ」になれるんだって



言い換え



機械翻訳



句読点の修正

\*T/N: "Todai" is short for "Tokyo Daigaku" (Tokyo University)



脚注の追加

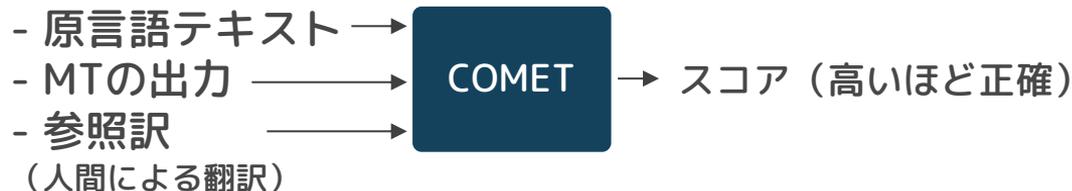
プロ翻訳者

## Q. この違いをどう評価する？

# ● 翻訳の評価手法について

## 2種類の評価手法

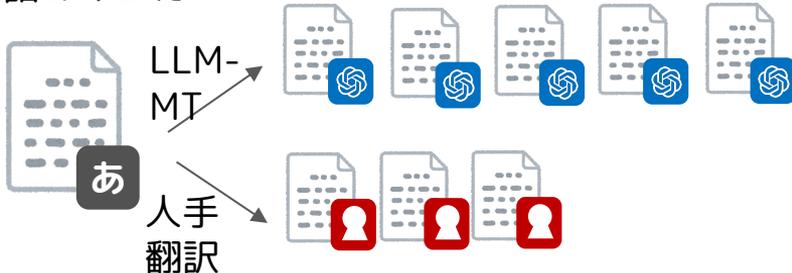
- 人手評価（プロの翻訳者による手作業）
  - 信頼性◎、高コスト、遅い -> 機械翻訳の開発現場で使いにくい
- 自動評価（正解データを用いた類似度計算）
  - 信頼性△～○、低コスト、速い -> 機械翻訳の開発現場で使いやすい
  - BLEU (2002)、COMET(2020)やその派生が広く使われてきた



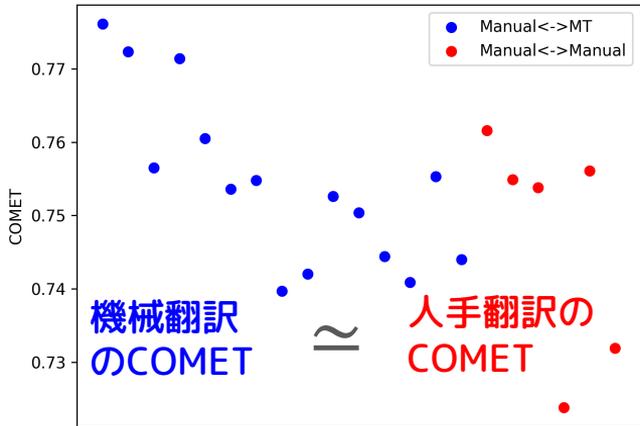
# 実験：自動評価 with マンガMT

原言語のマンガ

1. LLM-MT、人手でそれぞれ複数の翻訳版を作成



2. 様々な組み合わせでCOMETスコアを算出



課題  
自動評価手法COMETでは  
マンガ機械翻訳と人手翻訳の差を  
うまく評価できない…

※人手評価では、人手翻訳の方がLLM-MTよりも良いとされた



# ● Mantraのアプローチ

Confidential

1. 「良い翻訳」の要件を言語化する



2. 要件に基づき、人手評価を設計する



3. 人手評価のプロセスをLLMで自動化する



# 「良い翻訳」の言語化

Confidential

正確性：  
原文と似ていること

流暢性：  
訳文としての自然さ

作品としての面白さ

| Category | Subcategory           | Description   |
|----------|-----------------------|---|
| Accuracy | Mistranslation        | Translation is incorrect or inappropriate.  |
|          | Addition / Omission   | Information is added or omitted inappropriately.  |
|          | Inconsistency         | Inconsistent word choices, which makes it hard to read  |
| Fluency  | Unnatural word choice | Not grammatically incorrect, but feels weird to read (EX unusual collocation like "consume lunch" instead of "eat lunch").  |
|          | Unnatural structure   | Grammar is unnatural: passive/active, tense, etc. (EX: "I received a present from him." vs. "He gave me a present.") (EX: Japanese is written in the past tense, but it reads better in the present tense in English) |
|          | Spelling              | Spelling is wrong.  |
|          | Grammar               | Grammar is wrong.   |
|          | Punctuation           | Punctuation is wrong.   |
| Style    | Boring                | English could be perfect, but it is not fun to read and makes you not want to continue reading  |
|          | Genre fit / Tone      | Not boring, but the way in which it's written/made interesting doesn't match the work   |

# ● Mantraのアプローチ

1. 「良い翻訳」の要件を言語化する



2. 要件に基づき、人手評価を設計する



3. 人手評価のプロセスをLLMで自動化する

# ● マンガ翻訳の人手評価

## 💡 MQM (Multidimensional Quality Metrics) [Lommel+ 2013] に基づくマンガ評価

- マンガ翻訳経験の豊富なプロが原文と訳文を読み、文ごとに以下の分析を行う

1. 誤りの種類（前ページの観点）
2. 深刻さ：Critical/Major/Minorの三段階
  - Critical: 正しい理解を妨げるレベル
  - Major: 読みにくくなるが、意味はわかるレベル
  - Minor: 理解を妨げたり読みにくくしたりしないレベル
3. 誤りの詳細（自由記述）



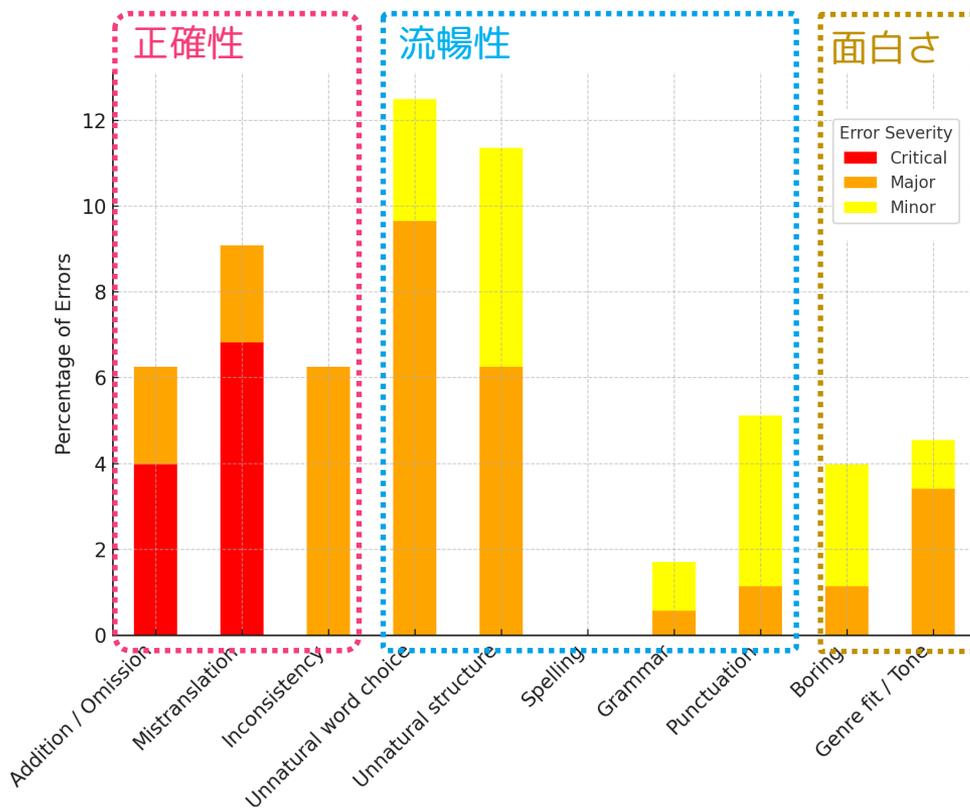
を読む



| E           | F              | G                     | H  |
|-------------|----------------|-----------------------|--|
| Page/Line # | Error Severity | Error Category        | Error Details  |
| 1/2         | Major          | Mistranslation        | Improperly identified subject/object. Does not read as a chapter title   |
| 2/2         | Minor          | Unnatural word choice | Overly literal translation that doesn't relate the statement to his current situation  |
| 3/1         | Minor          | Punctuation           | A question mark is needed here   |
| 4/16        | Major          | Unnatural word choice | Ruby text is ignored, translation sounds too unexciting  |
| 4/1-2       | Minor          | Unnatural structure   | Better flow between these two sentences would improve the translation  |
| 4/4         | Minor          | Grammar               | Second sentence is a fragment missing the subject. While this sort of construction is common in Japanese, it is not in English |
| 4/11        | Minor          | Punctuation           | Missing punctuation  |
| 4/12        | Critical       | Addition / Omission   | Translation omitted entirely   |
| 4/13        | Minor          | Punctuation           | Missing punctuation  |

誤りの分類・分析

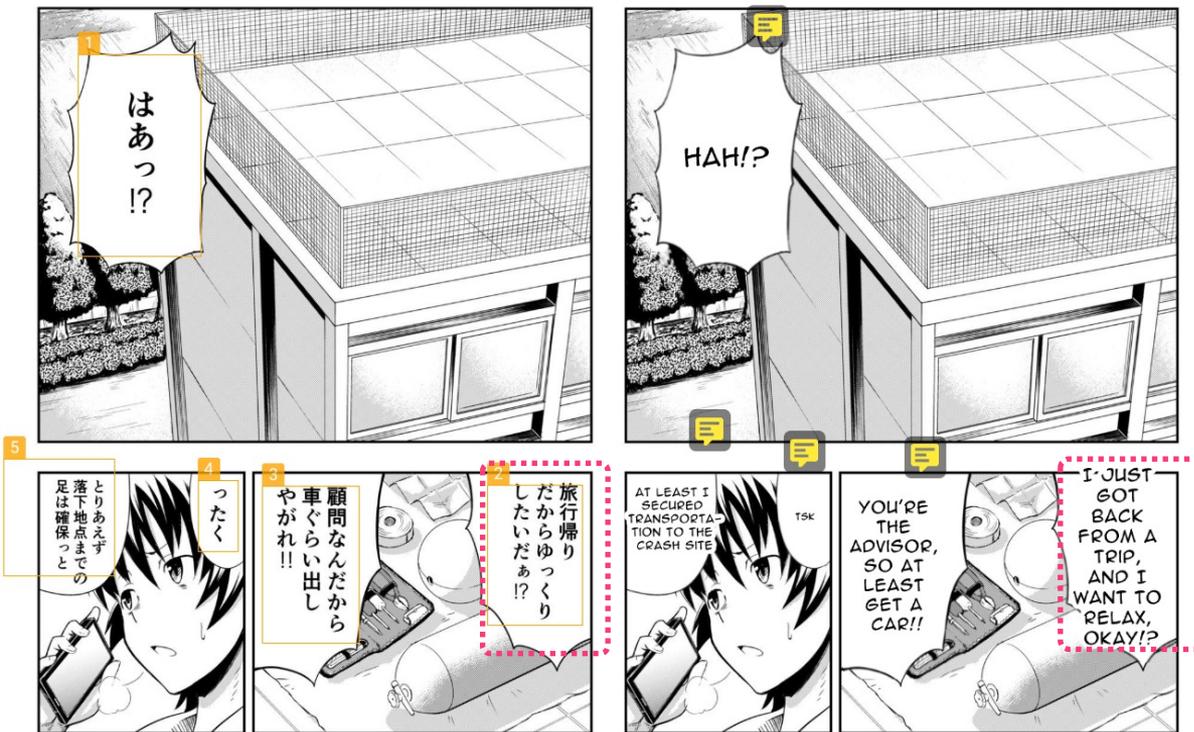
# ● 人手評価でわかるマンガMTの課題



- “Critical”なのは正確性に関するエラーで、全体の10%
- 最も多いエラーは不自然な用語や文構造  
(例：時制や能動・受動態)
- ジャンル・作品「らしい」翻訳にも課題あり

# エラー事例 主語目的語の推定ミス

💡 “Mistranslation”の6割は主語・目的語の翻訳誤り



主語の推定エラー

✗ I

○ You

# ● Mantraのアプローチ

1. 「良い翻訳」の要件を言語化する



2. 要件に基づき、人手評価を設計する



3. 人手評価のプロセスをLLMで自動化する

## 💡 人手評価の流れをLLMに模倣させる

プロンプトの構成：

・ 原文 / MT / 参照訳

・ 文脈情報

場面の説明、話者、聞き手、  
登場人物のプロフィール…

・ 評価タスクの指示

acceptableかどうか  
acceptableでは無い場合、エラーの詳細

・ 入出力のサンプル

人間による評価の事例

### Error Categories:

- **Mistranslation**
- **Definition**: The translation significantly distorts or alters the meaning of the original text.
- **Includes**: Misinterpretation of words, phrases, idioms, or context.
- **Excludes**: Minor differences that do not change the core meaning or acceptable alternative interpretations.
- ...

### Error Severity Definitions:

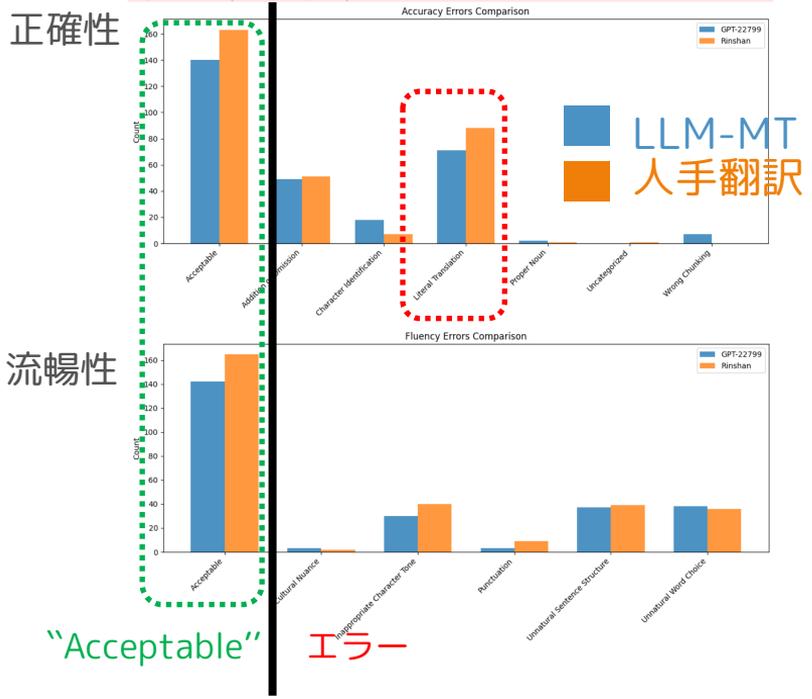
- **Critical**
- **Definition**: Errors that significantly distort the meaning or could mislead the reader.
- **Impact**: Misrepresentation of important information; could affect the storyline or character development.
- **Examples**:
  - Mistranslating a key plot point.
  - Omitting crucial dialogue that changes the understanding of the scene.
  - Misunderstanding of a subject or object.

### Output:

```
"line": 1,  
  "errors": [  
    {  
      "error_category": "Style and Tone",  
      "severity": "Minor",  
      "description": "'I'm sorry' is too formal and doesn't match ABC's  
interrupted speech after a minor mishap 'Excuse me' would better fit  
the context and maintain her character's tone."  
    }  
  ]  
}
```

# LLMによる自動評価の事例

## 💡 自動評価で 人手翻訳 と LLM-MT を比較



👍 正確性、流暢性ともに人手翻訳の方がMTより“Acceptable”判定が多い  
 →少なくとも人手翻訳とLLM-MTの分類機としては機能

### 今後の課題

- 😅 人手翻訳でも「直訳すぎる」エラーが多く指摘されている
- 😅 「面白さ」の評価は依然困難

# ● マンガ翻訳の難しさ

## 「面白い」翻訳

意味が伝わるだけではダメ  
作品として読んで面白い翻訳が必要

## 評価の難しさ

自動評価では人間と同じ精度  
プロ翻訳者の主観的な評価に依存

## 多様な文脈

前後の文、ストーリー、キャラ、  
文化的背景など多様な文脈が存在

## マルチモーダル

画像情報で訳し方が変わる  
キャラクターの口調や感情など

# ● マンガ翻訳における文脈の重要性

主語「私」の省略



文脈を考慮  
しないMT



人間による翻訳

I just wanted to...

単一の文が複数の吹き出しに分割

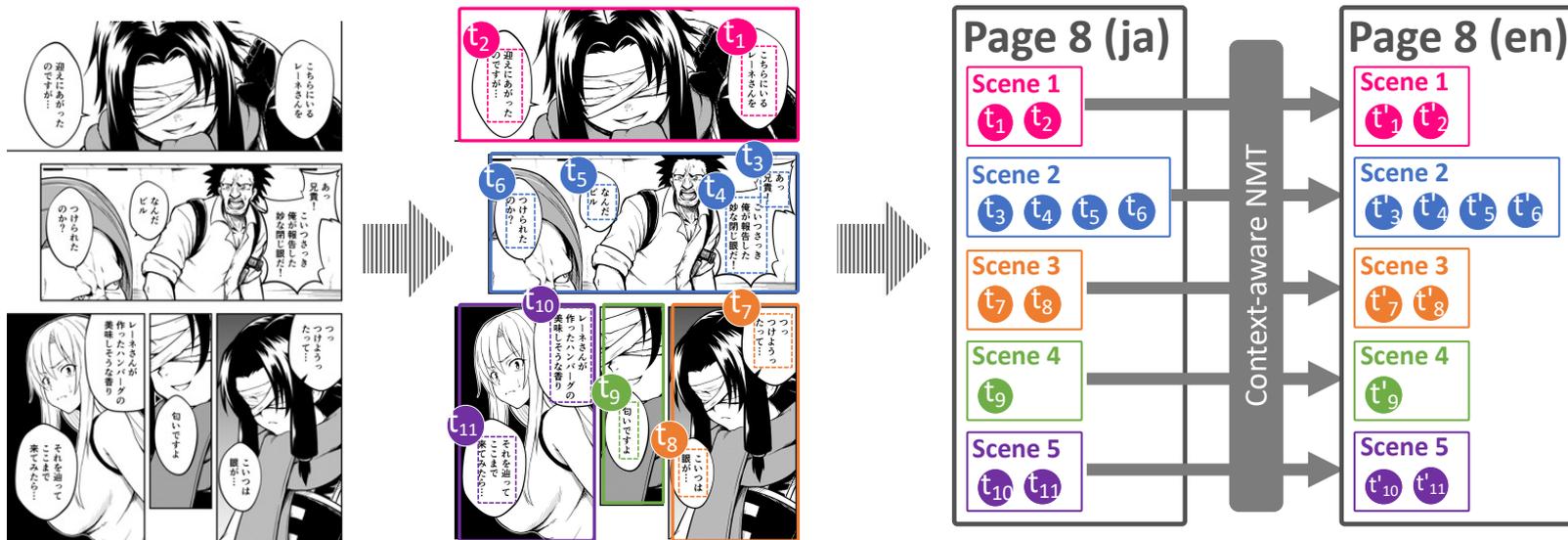
人間による翻訳

pick her up...

💡 マンガを正しく訳すには適切な文脈処理が必要

# 文脈情報を考慮するマンガNMT [Hinami+ AAAI2021]

💡 マンガの1コマを1つの「シーン」として捉え、  
場面に含まれるテキストを考慮すべき文脈とみなす



© 朽鷹みつき

# ● LLM時代のマンガMT

高性能なLLMにより、よりリッチな文脈情報を扱えるようになった

- ✓ 長い文脈；1コマから1巻へ
- ✓ 細かな翻訳ガイドライン  
話者、話者の性格、方言、スタイル、ジャンルの性質…
- ✓ 画像情報

# LLM時代のマンガMT | キャラ固有の翻訳



LLM



- 話者データベースなし
- ✓ファーストネーム
- ✓タメ口

話者データベース



LLM



- 話者データベースあり
- ✓苗字呼び
- ✓丁寧語

## ● まとめ

- マンガMTはマンガの世界展開を支える重要な技術
- LLMにより急速に進展したものの、課題も多い
  - 信頼できる自動評価
  - 「面白い」機械翻訳

現時点ではマンガ翻訳ではポストエディットは必須

# MANTRA

Delivering Manga across language barriers.