

## 職務経歴

### プリンシパル応用科学者

#### 医療データ分析部

#### Oracle

2024 年 7 月 - 現在

- 医療に特化した LLM の学習と医療系のエンティティリンクングに関する研究

### NLP リード

#### 開発チーム

#### Mantra 株式会社

2021 年 12 月 - 2024 年 5 月

- NLP の研究開発をリード（エンジニア 5 名）：
  - LLM を漫画翻訳に応用し、用語やストーリーを考慮した 3 万話以上の翻訳を達成
  - ノイズの多い OCR データを学習データから取り除き、単言語およびバイリンガルの LLM ファインチューニングを実施
  - Fairseq を用いて多言語かつ用語認識可能な翻訳モデルを訓練し、複数の言語での BLEU スコアを 10 点以上向上
  - プロの翻訳者と協力し、評価方法を改善
- Django と AWS を使用してアプリケーションのバックエンドを開発・保守

### NLP コンサルタント

#### フリーランス

#### 複数社

2020 年 1 月 - 現在

- 翻訳モデルが提案する単語から繰り返し選択して翻訳を構築できるインタラクティブな翻訳システムの開発をリード
- Google の「CANINE」論文の超効率トランスフォーマーを PyTorch で実装し、事前学習されたモデルを公開[\[link\]](#)
- spaCy を用いた情報抽出のパイプラインを書き直し、F 値のマイクロ平均を 20 点以上改善

### 研究実習生

#### 生産技術研究所

#### 東京大学

2020 年 12 月 - 2021 年 12 月

- 説明可能な検索拡張文法誤り訂正の手法の研究

### ソフトウェアエンジニア

#### 新商品開発部 & グローバル成長部

#### eBay

2016 年 11 月 - 2019 年 9 月

- eBay のチャットボットと中国版 iPhone アプリの自発的なメッセージを送信する Scala の分散スケジューラとその自動テストの開発を率い、1 人で保守
- Python で spaCy を利用し、商品レビューから有用な名詞と形容詞の組み合わせを抽出するブートストラップ自然言語処理を実装
- Ruby on Rails を利用し、非技術系社員が顧客へのメッセージをカスタマイズし送信をスケジュールできるウェブアプリケーションをゼロから開発

### ソフトウェアエンジニア

#### 国際配送管理部

#### eBay

2015 年 9 月 - 2016 年 11 月

- 提携先の運送会社からデータをインポートするパイプラインを再構築し、eBay の検索結果ページに載る国際配送料の精度を最大で 40% 改善
- Java と Spring で大規模（毎時数百万のリクエスト）なウェブサービスにおける保守と新機能の開発をし、JUnit と Mockito を利用してその自動テストも開発
- インフラやデータ整理などの手作業を自動化する取り組みを率い、国際出荷管理部チームの 1 週間当たりの手作業時間を数十時間減少させる事に成功

### インターン

#### 国際データインフラ部

#### eBay

2013 年 6 月 - 2014 年 9 月

- Hadoop での権限変更や運用タスクを自動化する Ruby on Rails ウェブアプリケーションの担当者となり、様々な新機能を開発

## 研究

### CoAM: Corpus of All-Type Multiword Expressions

Y Ide, J Tanner, A Noheji, J Hoffman, J Vasselli

ACL 2025 (Upcoming) [\[link\]](#)

### Context-Informed Machine Translation of Manga using Multimodal Large Language Models

P Lippmann, K Skublicki, J Tanner, S Ishiwatari, J Yang

COLING 2025 Oral [\[link\]](#)

### Project MOSLA: Recording Every Moment of Second Language Acquisition

M. Hagiwara, J Tanner

LREC-COLING 2024 Oral [\[link\]](#)

### Utilizing Longer Context than Speech Bubbles in Automated Manga Translation

H Kaino, S Sugihara, T Kajiwar, T Ninomiya, J Tanner and S Ishiwatari

LREC-COLING 2024 Poster [\[link\]](#)

### MWE as WSD: Solving Multiword Expression Identification with Word Sense Disambiguation

J Tanner, J Hoffman

EMNLP 2023 Findings & Poster [\[link\]](#)

### GrammarTagger: A Multilingual, Minimally-Supervised Grammar Profiler for Language Education

M Hagiwara, J Tanner, K Sakaguchi

Preprint 2021 [\[link\]](#)

## 受賞歴

- LREC-COLING 2024 Outstanding Paper Award

- 2018 eBay Critical Talent award (上位 1% の社員)

- 2016 eBay 社内イノベーション博覧会の発表者

## 学歴

### 計算言語学の理学修士

ワシントン大学 / 2021 年

### 計算機科学の理学士

エバグリーン州立大学 / 2015 年

### 交換留学

兵庫県立大学 / 2015 年

## 言語

英語 母語

日本語 日本語能力試験 N1