

bプロジェクト名:ニューラル言語モデルによる個人最適な日本語入力システムの開発
申請者名:三輪敬太、高橋直希

【提案プロジェクト詳細】

1 なにをつくるか

本提案書では「ニューラル言語モデルによる個人最適な日本語入力システムの開発」を主目標とし、ハイエンドのデスクトップ環境を対象として、新たな日本語入力システムの開発を提案する。これにより、多様なユーザーのための快適な日本語入力を開発し、公開する。

1.1 背景

日本語入力1億人を超える日本語話者がデジタル化社会で生活する上で、なくてはならないシステムである。我々は、現状の日本語入力の問題点として次の3点を指摘したい。

- 変換精度
- 個人最適化
- 予測入力

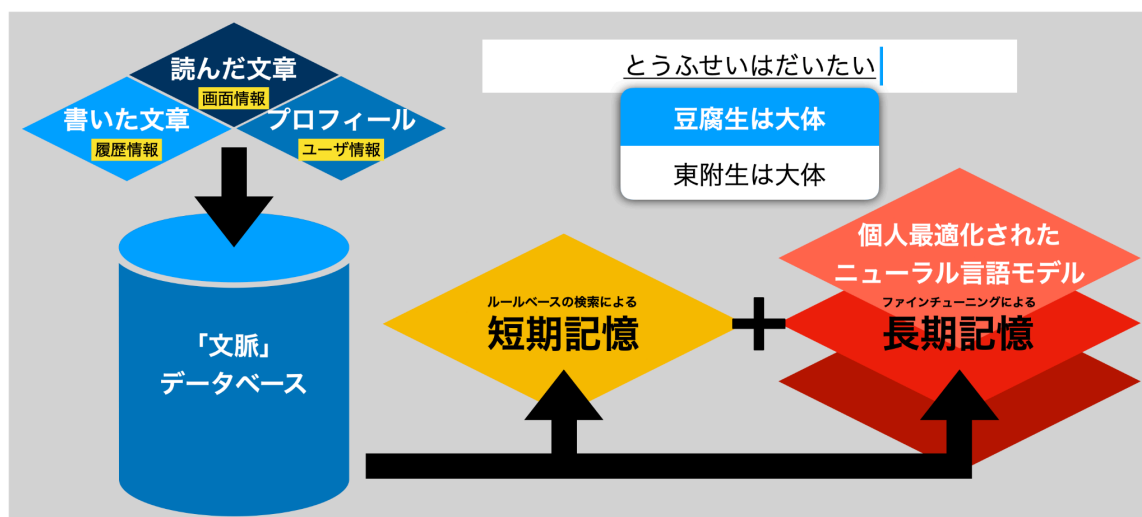
以下にその背景を述べる。

■変換精度 日本語入力には今なお十分に解決されていない精度の問題があり、不正確な候補への変換を優先することがある。多くのかな漢字変換システムは現在でも2-gramや3-gramなどの単純な言語モデルに基づいており、長距離の依存関係を考慮することが難しい。また、新語や流行語、俗語への対応が遅く、これらの言葉が入力できないことも多い。

■個人最適化 一方、ユーザーの持つ不満の多くは単なる精度の問題ではなく、パーソナライズをはじめとする広義の「文脈」理解への不満である。趣味、嗜好、年齢、職業、居住地などのプロフィール的情報や、「今どこにいるか」「なんのサイトを見ているのか」「最近読んだ漫画は何か」などのユーザーの状況など、システムが推薦すべきテキストは広い意味での「文脈」に依存する。従来のシステムはユーザー文脈の理解をシステムへのフィードバックにのみ頼っており、こうした個人最適化が不十分である。

■予測入力 近年は「予測変換」への不満の声もよく聞かれるようになった。この背景として、ChatGPTをはじめとする文章生成AIの普及がある。メールや説明文など、創造性の問われない文章を作成する際には、日本語入力システムが生成AIのように文章を自動生成し、入力を支援するのが望ましい。しかし現実の日本語入力システムは履歴に含まれる文や、入力途中の単語を予測変換する程度である。句や文、文章を予測入力するシステムは未だに実現されていない。

1.2 提案プロジェクト



提案プロジェクトでは、ユーザの広範な文脈情報を組み合わせ、高度に最適化した変換や予測入力を実現する日本語入力システムを開発する。このシステムは主に、ニューラル言語モデルの利点を活かしたかな漢字変換システムと、ユーザの文脈を記録するデータベースシステムからなる。

かな漢字変換システムとしてニューラル言語モデルを用いたアルゴリズムを提案する。このアルゴリズムは、数百語に及ぶ長いテキストを考慮することができ、従来のn-gram言語モデルベースのアルゴリズムと比べて「補償」と「保証」のような文脈依存の同音異義語における精度を向上させる。また、辞書データへの依存度が低いため、新語や流行語への対応にかかるメンテナンスコストを削減できるほか、個人最適化の観点でもメリットがある。実用のため、かな漢字変換のパフォーマンスを重視し、高速化・省メモリ化の方法も併せて提案する。次に、ユーザの「文脈」を記録するデータベースシステムを提案アルゴリズムと組み合わせることで、ユーザに最適化した変換と高精度な予測入力を実現する。変換に用いるニューラル言語モデルは、このデータベースを用いて定期的にローカル環境でファインチューニングされ、高度な個人最適化を可能にする。また、ファインチューニングでカバーしきれない短期間の文脈は文書検索の手法を用いて考慮に入れる。

最後に、提案する個人最適化に基づき、ユーザへの深い理解に基づいた予測入力システムを実現する。従来のシステムはユーザ理解に乏しいため弱い予測しかできなかったが、提案システムでは文単位の予測を含む高度な予測入力を実現する。

以上のとおり、提案プロジェクトは従来の日本語入力の課題である「変換精度」「個人最適化」「予測入力」の問題を解決し、新世代の日本語入力システムを実現する。以下に詳細を述べる。

1.2.1 ニューラル言語モデルを用いたかな漢字変換アルゴリズム

1.2.1.1 目的

すでに述べたものも含め、ニューラル言語モデルを用いたかな漢字変換には複数の利点がある。これらの利点は我々の実現したい柔軟な個人最適化の上で不可欠である。

- 長距離の文脈を正確に考慮することで、既存システムで誤変換を生じるケースをカバーできる
- ニューラル言語モデル自体をファインチューニングすることで、個人最適化をシンプルな形で、かつ効果的に実現できる(プロトタイプも参照)
- 多くの属性を持った辞書を明示的に利用しなくて良いため、開発プロセスがシンプルになる

一方で、日本語入力システムとニューラル言語モデルを組み合わせる方法には検討の必要がある。まず、従来の日本語入力では読みや頻度、品詞の属性を持った単語の辞書をベースとした手法が一般的だが、ニューラル言語モデルはこのような情報を明示的に持っていないため、既存のアルゴリズムが適用できない。また、予備実験として様々な事前学習モデルを用いて「タスクの説明→例題→問題」という構成(few-shot learning)によるかな漢字変換を行わせたところ、GPT-4を除いて入力に忠実な変換が全く出来なかったことがわかった。[ChatGPT-3が入力に忠実な変換に失敗する例](#)では「ようしょうき」を「要証紙(ようしょうし)」と変換してしまっている。この振る舞いは言語モデルがそもそも日本語の読みに関する情報をほとんど学習できていないことに由来すると考えられるが、本プロジェクトで実装するシステムではこの忠実性の問題を解決する必要がある。

1.2.1.2 アルゴリズム

高精度なニューラル言語モデルを活かした変換を実現するため、ニューラル言語モデルの利用を前提とした新規のかな漢字変換アルゴリズムを提案する。従来システム、few-shot learningによるシステム、提案システムの違いを表にまとめた。

	通常のシステム (n-gram)	ニューラル言語モデル +few-shot	ニューラル言語モデル +提案手法
精度	▲ 局所的な文脈活用	○ 長距離の文脈活用	○ 長距離の文脈活用
効率	○ 高速で省メモリ	✗ 低速でメモリ消費大	○ 投機的decodingと量子化で改善
忠実性	○ 辞書利用で忠実	✗ 入力を無視	○ 読み制約で担保
学習	▲ 履歴ベース	▲ プロンプトで制御	○ 文脈データベースに基づき個人最適化

提案手法では、次トークン予測を繰り返しながら文を生成するビームサーチのような手法を、適切な追加制約(読み制約)に基づいて利用することでかな漢字変換を実現する。

■精度 ニューラル言語モデルが意味関係を詳細に考慮できることから、大幅な改善が得られる。後述のプロトタイプでは既存の変換器に比した性能向上を確認した。

■効率 次節1.2.1.3で詳しく議論する。

■忠実性 これを担保するため、読みによる制約を用いる。例えば「紙」を「き」とは読まないことから、「ようしょうき」に対して「要証紙」との変換が不適切であることを検出し、生成を防止する。従来のシステムは単語単位に基づいたシステムであり、「ようしょうき→幼少期」のような対応関係を示す辞書データから変換していたが、我々のシステムはこれを「よう→幼」「しょう→少」「き→期」と漢字1文字単位で与える。単語単位は探索する状態数を減らす上で効率的であり、文字単位では「よう→幼」「しょう→賞」のような通常起こり得ない熟語を検討する必要が発生してしまう。しかし、ニューラル言語モデルとビームサーチにより、これを効率的に避けることができる。

■学習 既存の辞書ベース手法は単語単位を利用するが、こうした手法では「ふかばる→深掘る」などの新語への対応が困難である。これでは個人最適化の観点で必要な柔軟性を実現できない。また、読みと単語の対応関係の保証も簡単ではなく¹、辞書データのメンテナンスが難しい。提案アルゴリズムは漢字1文字単位の読み制約を基本とした緩い制約のみを用いることで、こうした柔軟性を確保する。

このような形で、シンプルなニューラル言語モデルを利用しつつもかな漢字変換エンジンとして実用可能なシステムを構築する。

1.2.1.3 省メモリ化・高速化

日本語入力は重要だが、他のアプリケーションと共用する補助的なシステムである。このため、システムのリソースを過剰に消費してはならない。我々は省メモリ化と高速化の両面からシステムの効率化を図る。

省メモリ化については、言語モデルの量子化により言語モデルのメモリ消費を抑える。プロトタイプ(後述)ではGPT-2モデルに対して`llama.cpp`による量子化を適用し、単純な生成時のメモリ消費を150MB程度に抑えた。これはAppleの「日本語入力プログラム」の200MB程度よりも小さい。

高速化については、投機的デコーディングと呼ばれる手法を採用する。これは高速で低精度な方法で先読みして候補の生成を行い、高精度で大きな言語モデルによってその妥当性を検証しながら生成を進める高速化手法である。先読み候補が一致した場合に高速化が得られる。かな漢字変換では、最長一致法のようなシンプルなアルゴリズムでもある程度の精度が得られることが知られているため、このような高速で低精度のヒューリスティクスを先読みに用いることで、大幅な高速化が期待できる。

1.2.2 「文脈」のデータベースシステム

1.2.2.1 データベースの構築

上記のニューラルかな漢字変換システムとともに高度な個人最適化を実現するために、我々は「文脈」のデータベースシステムを開発する。これはユーザ自身が自身のプロフィールや、過去に書いたテキスト、ブラウザの閲覧履歴、画面上に表示されたテキストなど、最適化に有用と考えられる情報を任意で提供し、それを日本語入力のために効率的な形で記録するシステムである。

具体的には、以下のような情報をデータベース化する。

■ユーザのプロフィール ニューラル言語モデルはプロンプトによって挙動を大きく変えることができる。そこでユーザのプロフィールをプロンプトに含めながら先述の変換アルゴリズムを利用することで、低コストの個人最適化を実現できる。

■ユーザが書いたテキスト 多くの場合、ユーザはSNSやチャット、メモアプリなど、様々な場面ですでに多くのテキストを生産している。これらをユーザが任意でシステムに与えることで、過去のユーザのテキストに基づいてニューラル言語モデルをファインチューニングすることができる。

■ユーザが読んだテキスト 多くのユーザは全くの造語ではなく、過去に自身が読んだテキストから引用する形で、新しい言葉を使う。SNSで「雷雪」という珍しい気象現象を知り、それについて言及する、というような振る舞いはその例になる。これらのテキストを日本語入力システムが知ることができれば、さらに高度な文脈の把握が可能になり、それに基づいた変換も可能となる。ユーザがこのようなテキストデータをシステムに与える方法として、例えば以下のようなシステムをそれぞれ追加で利用できるようにする。

- ブラウザ拡張機能による自動記録
- ユーザが提供したテキストファイルやPDFファイルからの自動記録
- ユーザが許可する範囲の画面情報からOCRによって自動記録

¹ 提案者は最近AppleとGoogleの日本語入力エンジンに共通して「しんこうほうほう→進行方向」という誤ったデータが含まれていることを発見したが、元を辿るとこれはWikipediaに「進行方向(しんこうほうほう)」と書かれていることが原因であった。辞書の管理は大企業であっても困難なタスクである。

1.2.2.2 日本語入力システムとの統合

記録した文脈データと日本語入力システムの統合は「長期記憶」と「短期記憶」の2段階で行う。

「長期記憶」では、ユーザの文脈に沿ったデータでニューラル言語モデルをファインチューニングすることで、より個人最適化された変換を実現する。ファインチューニングは数日に一度、ユーザがPCを使わない時間帯に自動実行される。これはいわば「ユーザ専用モデル」の学習である。

「短期記憶」は、長期記憶がカバーできない新しい情報を高速に変換に反映するために用いる。ブラウザ拡張や画面情報、入力中のテキストデータなどを対象とし、接頭辞検索に優れたデータ構造でデータを整理し、変換時に検索をかけ、合致するものがあれば優先的に候補に加えることで、短期的な文脈を考慮した変換を実現する。検索拡張生成(RAG)や、コンテキスト幅の広い言語モデルを用いた文脈内学習(In-Context Learning)も検討するが、これらの手法では速度面で変換に十分なパフォーマンスが実現できない可能性がある。

これらの統合により、ユーザの文脈に深く適応した日本語入力システムを実現する。

1.2.3 予測入力システム

n-gramをベースとする従来型の日本語入力では、予測処理はあくまで変換の拡張にとどまっていた。一方、ニューラル言語モデルを基盤とする提案システムはより長い予測入力を容易に実現できる。

プログラマ向けのコード生成ツールであるGitHub Copilotは入力途中のプログラムの続きを1行～数十行に渡り生成し、インラインで提案する仕組みを持つ。Apple PlatformやWindowsなどでも英語を中心に類似のインラインの予測入力機能がサポートされつつあるが、これらはただ数単語の予測にとどまっている。我々のシステムはユーザの文脈に対する強いアクセスを持つことを活かし、GitHub Copilotのような長文の提案を日本語入力と統合して実現する。有用な提案システムに対してユーザはある程度「待てる」ため、RAGを含むある程度の処理時間の必要なシステムを構築することもできる。

1.3 プロトタイプ

プロジェクトの提案にあたり、実現可能性を示すため「ファインチューニングによって個人最適化を行った言語モデルに基づくかな漢字変換改善」をテーマとするプロトタイプを開発した。

システム構成

提案システムは次の2要素からなる。ニューラル言語モデルの準備は高橋が、かな漢字変換エンジンの準備は三輪が主に担当した。

■ファインチューニングしたニューラル言語モデル 日本語GPT-2モデル[rinna/japanese-gpt2-medium](https://huggingface.co/rinna/japanese-gpt2-medium)を基盤モデルとし、高橋の約2万件のツイート履歴によってファインチューニングした。この言語モデルは提案システムにおける「文脈のデータベース」に基づいて構築される言語モデルを簡易的に構築したものである。

■ニューラル言語モデルを統合したかな漢字変換エンジン 三輪開発の[AzooKeyKanaKanjiConverter](https://github.com/azookey/azookey-kanakana-converter)をベース実装とし、候補のランキング処理にニューラル言語モデルが出力するテキストの確率に基づいた並び替え処理を追加した拡張システムを開発した。この変換エンジンは提案システムが目指す「ニューラル言語モデルを前提とした変換アルゴリズム」とは異なるが、ニューラル言語モデルの長所をある程度反映できる仕組みになっており、短期間で開発するプロトタイプでは妥当だと判断した。

システムの振る舞い

提案システムは高橋に個人最適化された挙動を示すことが期待される。そこで高橋の個人的な傾向として「数学」「クレヨンしんちゃん」「出身高校」を例とし、素のAzooKeyKanaKanjiConverterの変換結果(GPT-2なし)、ファインチューニングを行っていないGPT-2モデルを用いた場合の変換結果(GPT-2 Rinna)、ファインチューニングを行った場合の変換結果(GPT-2 高橋化モデル)を比較した。

入力	GPT-2無し	GPT-2 (Rinna)	GPT-2 (高橋化モデル)	目標候補	成功
すうがくのかん	数学の漢	数学の漢	数学の環	数学の環	YES
るべーぐそくど	ルベーク速度	ルベーク速度	ルベーク測度	ルベーク測度	YES

のはらしんのすけ	野原慎之助	野原慎之助	野原しんのすけ	野原しんのすけ	YES
のはらひろし	野原広	野原浩	野原浩	野原ひろし	NO
とうふせい	豆腐性	当不正	豆腐生	豆腐生（東工大附属高校＝豆腐）	YES

失敗するケースもありつつ、概ね個人最適化が実現できていることがわかる。変換精度では、AzooKeyKanaKanjiConverterの持つテストケースにおいてGPT-2 Rinnaを追加することによる5~10%の正答数増加が確認され、個人最適化版モデルでも同程度の性能向上が保たれていた。プロトタイプについては詳細な分析を進めているが、一旦の結果として、ファインチューニングに基づく言語モデルの個人最適化が実現可能であり、これをかな漢字変換システムに反映することも可能であることを示すことができたと考えている。

1.4 開発ステップ

フェーズ1: 基盤構築

フェーズ1では提案システムの中核となるシステム群を構築する。

文脈データベースシステム

先に述べた通り、ユーザが任意に提供したデータを元にデータベースを構築するシステムを実装する。データベースはSQLiteをベースとする。また、[Llamaindex](#)などのRAG用途のデータベースシステムについてもこの段階で利用可能性を探る。文脈取得システムとして、ユーザが選択したテキストファイル等をデータベースに統合するシステムや、Chrome拡張による閲覧サイトの自動記録システムを作る。

かな漢字変換システム

先に述べた通り、ニューラル言語モデルを前提としたアルゴリズムを実装する。かな漢字変換システムは変換アルゴリズム以外にも様々なコンポーネントを含む複合システムであるため、変換アルゴリズム以外の部分は既存のOSS実装であるAzooKeyKanaKanjiConverterや[Mozc](#)を参考にする。最低限実用可能なものを目指し、高精度化・高効率化はフェーズ2以降で行う。

ユーザーインターフェース

提案システムはユーザが利用可能なデスクトップIMEプロダクトの開発を目指す。このため、フェーズ1の段階からmacOS向けに実用的なGUIを構築する。macOSデスクトップIMEの開発は複雑な箇所が多いため、既存のOSS実装であるMozc、[AquaSKK](#)、[EmojiIM](#)などを参考にする予定である。

フェーズ2: 主要機能開発

フェーズ2では、各システムの主要機能を開発する。

文脈取得システム

文脈データベースシステムをユーザが利用しやすくするため、ユーザのパーミッションのもと、OCRによってPCの画面に表示されている情報を取得し、データベースに記録するシステムを構築する。これにより、より幅広いユーザ文脈への暗黙のアクセスが可能になる。加えて、取得したテキストの妥当性を検証し、OCR固有の誤りを避けながら学習する方法を検討する。

かな漢字変換システム・予測入力システム

フェーズ2では変換アルゴリズムを改善し、個人最適化を行っていない状態で既存の変換エンジンと同等以上の変換精度と、ハイエンドのPC端末においてストレスのない変換速度の実現を目指す。フェーズ1の結果に応じてある程度進め方は変わるが、提案時点では以下の取り組みを計画している。

1. 量子化モデルを用いて変換の省メモリ化を図る
2. 最長一致法を利用した投機的デコーディングの実装を行い、変換の高速化を図る
3. GPT-2で決め打ちしていた言語モデルをより広い範囲で検討し、適切なモデル選定を実施する
4. 日本語の読み情報をより理解する言語モデルをファインチューニングなどによって得ることで、これを用いた実用レベルの精度を実現する

また、過去の入力などの取得情報と直前の文脈に基づき、1行以上の長い文章を提案する予測入力エンジンを開発する。

フェーズ3: 文脈データベースの多様化と入力システムの高機能化

フェーズ3では文脈データベースがソースとできる外部情報を多様化する。また、日本語入力システムとしては高機能化を進め、既存の日本語入力システムからの移行が可能な実用プロダクトを目指す。

文脈取得システム

近年普及が進んでいるパススルーを持ったヘッドセット端末を通じて、ユーザと同等の視覚を用いた文脈把握を実現する。パススルー式のヘッドセット端末の画面全体の表示情報や、外部カメラ・マイクの音声情報やその他の情報(位置情報・気候情報など)を取得して文脈データベースに記録するシステムを構築する。

かな漢字変換システム・予測入力システム

実用レベルの日本語入力システムに必要な機能を実装するとともに、さらなる高速化や高精度化を図る。今回提案する制約付き生成では、読み制約を緩和することによって誤入力訂正や曖昧入力に相当する振る舞いを容易に実現できる。また、今回提案する制約付き生成では、読みを追加する形でユーザ辞書を効率的に導入できる。

予測入力システムでは、RAGなどの高度な処理を検討し、さらに流暢な入力を実現する。

今回提案するシステムは潜在的に中国語などを含むより多くの言語に一般化できる可能性があるため、この領域も実験的に検討する。

プロダクト化

GUIツールやインストーラ、自動更新機能、ドキュメントなどを整備し、非開発者のユーザが容易に導入・利用可能な状態を目指す。

2 どんな出し方を考えてるか

日本語入力は、どんなに優れたPoCを実現したとしても利用するユーザがいなければ存在価値がない。提案プロジェクトの成果物はハイエンド環境で実際に利用可能なプロダクトとすることを目標として開発する。提案システムは無償で全てのユーザに提供し、開発者向けにオープンソースで公開する。

公開にあたってはmacOSを主な対象としたアプリケーションとするが、最終的にはクロスプラットフォームで実用的に動作することを見据えて開発する。

3 斬新さの主張、期待される効果など

3.1 斬新さの主張

提案システムは、ユーザのフィードバックを中心とした最適化を行う既存システムとは異なり、ユーザの持つ文脈にアクセスし、ユーザと同等の知識を持ったデータベースを利用することで、高度に個人最適化された日本語入力を提供する点で革新的である。文脈にアクセスするシステムはATOKインサイト²やKukura³などで試みられているほか、位置情報などに基づいたシステムもiWnn⁴などでは実現されているものの、提案システムのようにアクセスした情報に基づいてニューラル言語モデル自体のファインチューニングまでも実施する方式は類を見ない。ニューラル言語モデルの柔軟性は無制限のユーザへの最適化を実現できる。また、画面情報へのアクセスなど、文脈アクセスのためのシステムを多様化し、できる限り暗黙にすることで、ユーザの注意を奪わないよう工夫している。ニューラル言語モデルを活用した制約付き生成によるかな漢字変換アルゴリズムと、ユーザの文脈を広範に学習する言語モデルを組み合わせることで、日本語入力におけるユーザ最適化を根本から改善することができる。

² [\[057984\]参照している文章から候補を推測して表示するー時文書学習「ATOKインサイト」ー](#)

³ [予測入力の拡張](#)

⁴ https://socialsolution.omron.com/software/products/product_text/iwnn/

3.2 期待される効果

3.2.1 高度な個人最適化

提案システムは高度な個人最適化を実現し、専門用語や表現の入力を強力にサポートする。専門用語の利用という一部の専門家に限られた話に見えるが、漫画の登場人物名、ゲームのアイテム、スポーツ選手、学校の流行り言葉など、限られた分野でのみ利用される言葉は存外に多い。従来の「辞書」という固定的な語彙リストに基づいたシステムとは異なり、ニューラル言語モデルと文脈データベースシステムに基づいた個人最適化は柔軟に新たな用語や表現を学習し、より幅広い分野の日本語入力をより流暢にする。

また、提案システムがユーザの文脈により多くのアクセスを持つことにより、自然で有用かつ長い文の予測入力が可能となる。ChatGPTなどに頼った文章生成では、適切な出力を得るためにユーザが明示的に文脈を説明する必要がある。元から文脈を適切に把握した提案システムにより、アプリケーションを切り替えることなくこうした処理を行えるようになるのはユーザにとって大きな改善である。これにより、単なる入力の効率化のみならず、ユーザの創造性を刺激し、発想を豊かにする効果も期待できる。

一般的な利用者のみならず、地域方言をはじめとする言語的マイノリティにも最適化された入力体験を提供すると期待できる。これにより、言語的背景によって入力効率に差が生じるという不平等が解消され、誰もがストレスなく自らの言葉で表現できる環境が整う。この言語的平等の実現は、言語的な発信の障壁を減らし、より活発な情報発信を促進する。多様な背景を持つ人々のコミュニケーションを促進し、新たなアイデアが交錯する創造的な社会の形成に寄与する。地域の文化の保存・振興や、学術分野の発展にもつながる可能性を秘めている。

3.2.2 新世代の日本語入力

日本語入力ソフトウェアの開発では、競合製品とのアラインメントが重要である。文脈データベースに基づいた高度な個人最適化システムを持った我々の日本語入力システムがユーザに利用され、常識となることにより、様々な日本語入力ソフトウェアがこのような新標準に対応していくことが考えられる。近年の日本語入力は安定期を迎え、日本語入力に関心を持つ開発者も減り、革新的な変化というよりは順当な改善が続いたように感じられるが、提案プロジェクトを嚆矢とし、日本語入力への関心が再び高まり、新技術の提案や開発が盛り上がることを期待している。

4 具体的な進め方と予算

4.1 開発

開発場所

各々の自宅

使用する計算機

本プロジェクトにおいてはニューラル言語モデルなどの構築などにおいて計算リソースが要求される。これについては変換エンジンの開発を主として三輪の所持する計算リソースにより開発を行うが、随時AWSなども利用する。そのほかは以下の個人所有の計算リソースを利用する。

- Macbook Air, M1 2020, macOS Sonoma 14.2.1(高橋)
- Mac mini, M2 Pro 2023, macOS Sonoma 14.2.1(三輪)

プログラミング言語

- JavaScript / TypeScript, HTML, CSS(Chrome拡張)
- Swift(macOSアプリケーション、変換エンジン)
- Python(言語モデルの学習等)

使用ツール

- Slack, Notion, Google Drive, GitHub, VS Code, Xcode, Gitなど

作業分担

以下を目安とし、適宜分担する。

- 三輪: ニューラル言語モデルを用いたかな漢字変換アルゴリズムの実装、変換エンジンの高速化・高精度化、および予測入力システムなどの開発
- 高橋: ユーザーインターフェース、文脈データベースシステム、文脈取得システム(Chrome拡張、画面

情報OCR)の開発

開発線表

2024年3月～5月: プロトタイプ - 概念実証(完了)

高橋: ファインチューニングを行った言語モデルの準備

三輪: ニューラル言語モデルを用いたかな漢字変換エンジンのプロトタイプ開発

6月～7月: フェーズ1(基盤構築)

高橋: 文脈データベースの基本システム構築、長期記憶システムの構築、macOSアプリ化

三輪: ニューラル言語モデルを用いたかな漢字変換アルゴリズムの実装、短期記憶システム構築

8月: フェーズ1の統合・検証

高橋・三輪: フェーズ1で開発したシステムの統合、変換精度の検証と改善、方針確認

9月～10月: フェーズ2(機能開発)

高橋: OCRによる画面表示情報の取得システムの構築・検証

三輪: 予測入力システムの実装、変換エンジンの効率化、ニューラル言語モデルの改善

11月: フェーズ2の統合・検証

高橋・三輪: フェーズ2で開発した機能の統合、変換精度の検証と改善、方針確認

12月～2025年1月: フェーズ3(多様化と高機能化)

高橋: 位置情報や音声情報、VRヘッドセットに基づく周辺情報などの取得システムの追加

三輪: さらなる高速化・高精度化、ユーザ辞書などの高機能化、検索拡張生成による予測入力の検討

高橋・三輪: フェーズ3で開発した機能の統合、プロダクト全体の完成度向上

2月: ユーザ向けプロダクトとしての仕上げ・報告準備

高橋・三輪: GUI、インストーラ、アップデート機能などの整備、プロダクトとしての最終調整

高橋・三輪: 未踏最終報告の準備、プロジェクトの総括

開発にかかわる時間帯と時間数

平均して一人当たり週19時間程度の活動を行う。日中は大学などの活動があるため、平日午後・土日を主な活動時間とする。月80時間/人 x 9ヶ月より1440時間の活動を予定する。

4.2 予算

予算内訳は以下のとおり。

活動時間	活動時間 x 人件費
1440時間	2,880,000円

5 提案者(たち)の腕前を証明できるもの

三輪

実績

日本語入力およびそのGUIを中心に、NLPやソフトウェアエンジニアリングの分野で実績がある。iOSの日本語入力キーボードアプリとして「[IazooKey](#)」の開発を2020年9月に個人で開始し、[Google日本語入力 / Mozc](#)を参考に、変換エンジン、辞書データとその生成システム、GUIをそれぞれ独自に構築した。開発開始から3ヶ月で通常の変換や予測変換、曖昧入力などの機能を完成させ、[App Storeで公開](#)。その後過去3年半以上継続的に開発している。2023年2月には[ソースコードをGitHubで公開](#)し、OSS化を果たした。GUI面ではキーボード配列のカスタマイズの機能を搭載するなど、痒いところに手の届くアプリケーションとして徐々に知名度を高めており、[日経産業新聞でも紹介](#)された。[App Storeでは公式キュレーションに入ったこと](#)もあり、執筆時点でApp Storeで平均評価4.6 (レビュー数160)と1万DLを達成した。GitHubでは200以上のスターを獲得している。[変換エンジン](#)

は独立のライブラリとして公開しており、複数の第三者の製品でも利用されている⁵。また、iOSの国内最大規模のカンファレンスイベントである[iOS Developer Conference Japan 2023](#)で登壇し、日本語入力開発の知見を発表した。

その他の活動

卒業研究 / [大関研究室](#) (東京大学)

大学では卒業研究として「人間の視線データに基づいて言語モデルの振る舞いを評価する」という枠組みでの研究に取り組んだ。2024年3月の自然言語処理学会ではこの結果を元に、[GPT-2の振る舞いをトークナイザの違いに関して評価した研究を発表](#)した。

インターン / [Turing株式会社](#)

創業初期の2021年より2024年現在までインターン。初期の自動運転デモにおいて自動運転を担うAIの開発で中心的に貢献したほか、自動車の運転の制御、CANによる通信の解析、信号機認識AIの開発、車載ディスプレイのAndroidベースのOS開発およびJetpack Composeを用いたGUI開発など、ソフトウェア開発と機械学習を中心に開発に参加。共著だが成果の一部は論文にもなっている^{6,7}。

インターン / [Google Japan \(Mozc・Google日本語入力・Gboardチーム\)](#)

azooKeyの開発をきっかけとして、2022年より二度インターン。[Mozc](#)の開発に参加し、異なるフォント環境で変換候補を最適化する実装に取り組んだほか、辞書データの検証、変換の高速化、特殊文字の追加など様々な面で貢献。クローズドソースのため詳細は伏せるが、Android版の[Gboard](#)の開発に参加。変体仮名のオープンソースフォント開発プロジェクトである[Noto Hentaigana](#)にも社内でも貢献した。

高橋

実績

2019年アプリ甲子園においてARナビゲーションアプリ「HybridMap」で全国3位を受賞。韓国の釜山で行われた[韓国科学アカデミー科学フェア2019](#)で「剣道の自動審判システム」で優秀革新的研究賞受賞。3年次課題研究にて第63回日本学生科学賞入選1等受賞および[第4回関東・甲信越静地区高校生探究学習発表会 ポスター日本語部門優秀賞](#)。国際イノベーションコンテストにて3位入賞。2022年、企画開発をした「おしゃべりひろゆきメーカー」が[ACC TOKYO CREATIVITY AWARDS](#)にて、総務大臣賞/ACCグランプリ、ACCシルバーを2部門で受賞。2023年Apple主催の[WWDC Swift Student Challenge](#)でWinner1に選出。また、2024年にazooKeyの漢字変換エンジンを利用したOSSの[Apple Vision Pro用日本語エディタ「azooEditor」](#)を開発しApp Storeでリリース。Apple Vision Proの日本語利用者より多くの喜びの声を受けている。またazooKeyの開発にも貢献した。

就業経験

[株式会社LetterFan](#)

2020年に著名人・アイドルとのビデオレターアプリ「LetterFan」のカメラアプリの実装および改善施策の実施

[株式会社CoeFont](#)

2020年に創業メンバーとして創業。株式会社CoeFontにてセールスおよびカスタマーサポート、サービス開発およびDX業務を行う。現在は同社のマーケティングとして「おしゃべりひろゆきメーカー」などの企画・開発を担当。

その他

詳細についてはNDAにより記載できないがiOSキーボードの開発などを受託し開発。

6 プロジェクト遂行にあたっての特記事項

三輪: 未踏期間中に所属することになる大関研究室(東京大学)と勤務先の了解は得ている。

高橋: 特になし

⁵[「Forethumb」](#)「[推し活アプリ オシバナ\(Oshibana\)ウィジェット](#)」「[azooEditor](#)」など

⁶[SuperDriverAI: Towards Design and Implementation for End-to-End Learning-Based Autonomous Driving](#)

⁷[自動運転のための大規模走行データセットを用いた深層学習による信号機認識](#)

7 IT以外の勉強、特技、生活、趣味など

三輪

東京大学教養学部後期課程所属。言葉に広い関心を持ち、計算心理言語学を学んでいる。日本語全般に興味があり、日本語の用例収集が趣味である。最近の関心は「深掘り」「片思い」などの動詞が「深掘り」「片思い」などから新造される現象である。また日本の古字である「変体仮名」が好きで、出先で見かけた変体仮名の用例の写真を撮って集めている。書体やタイポグラフィにも関心があり、オリジナルの書体をデザインすることもある。

高橋

現在は早稲田大学基幹理工学部数学科に所属しており、集合論に興味がある。研究室配属などはまだだが、今後は集合論を中心として研究をしていきたいと考えている。また最近マーケティングについて学んでおり、現在は効率的な広告運用やTikTokやYouTubeなどのShort Videoコンテンツの効率的なバイラルなどを個人的に試行錯誤して研究している。またプレゼンテーションなどが好きでWIPO Show and Tellプレゼンテーションコンテストなどにも開発したアプリのプレゼンなどを行い、1位獲得。

8 将来のITについて思うこと・期すること

三輪

現代社会には言語の壁を個々人の努力による研鑽で突破しようとする/させようとする風潮がある。しかし、このようなやり方には当然に限界があるし、膨大な時間が研鑽に注がれることによる損失は大きく、全体最適化を妨げている。近年著しく発展した自然言語処理技術をはじめとするIT技術に基づく仕組みとシステムによってこうした風潮を打破し、誰もが言語的なアクセシビリティの保障を享受する社会が実現することを強く期待している。

高橋

プログラミングスクールでの指導経験から、ChatGPTなどの補助ツールの登場にもかかわらず、多くの人にとって未だにプログラミングは難しいものである。プログラミングの最大の障壁は、自分の作りたいものを言語化し、それを詳細な要件にブレークダウンすることだと考える。将来的には、こうした障壁を下げるツールの発展により、より多くの人々がプログラミングに取り組める世界が実現することを切に願っている。同年代の友人にもプログラミングに興味を持ってもらい、共に学び合える環境が広がればと思う。