

2025年3月4日

株式会社博報堂DYホールディングス

株式会社博報堂テクノロジーズ

博報堂DYホールディングス、博報堂テクノロジーズ、 言語処理学会第31回年次大会(NLP2025)にプラチナスポンサーとして協賛

株式会社博報堂DYホールディングス（東京都港区、代表取締役社長：水島正幸、以下 博報堂DYホールディングス）と株式会社博報堂テクノロジーズ（東京都港区、代表取締役社長：米谷 修）は、2025年3月10日（月）～14日（金）に長崎で開催される言語処理学会第31回年次大会（NLP2025）にそれぞれプラチナスポンサーとして協賛し、現地でのスポンサーブースの設置や、発表論文へのスポンサー賞の授与を実施いたします。また、研究成果について口頭発表1件、ポスター発表2件、ワークショップ発表1件を実施いたします。

言語処理学会全国大会は、日本の言語処理の研究成果発表の場および国際的な研究交流の場として1994年に設立された言語処理学会が年次で開催している学術発表イベントです。昨年開催された第30回年次大会では、総参加者数2121名、総発表論文数は559件にのぼるなど、自然言語処理領域における国内最大の研究発表の場として、大学、企業を問わず国内の研究者の注目を集めています。

博報堂DYホールディングスは、AI領域における先端研究と技術開発を推進するために、Human-Centered AI Institute（代表：森正弥 以下、HCAI）を設立し、生活者発想に基づいた「人間中心のAI活用」と「人間の創造性の進化・拡張」を実現するための研究開発に取り組んでいます。

また、博報堂テクノロジーズは、プロダクト開発センターを中心に広告におけるAIの研究開発からプロダクトの実装までを行っています。AIのエンジン部分から生成AIをいかしたRAG、AIエージェント等の各種アプリケーションまでフルスクラッチで開発しており、広告業界に対しユニークなプロダクト・技術を打ち出しています。

この度の協賛は、国内の自然言語処理領域の研究コミュニティへの支援を目的とするものです。今後も、博報堂DYホールディングスおよび博報堂テクノロジーズは、HCAIやプロダクト開発センターの活動を中心に、博報堂DYグループのけん引役として、幅広い研究開発活動を通して革新的なAIソリューションの開発を目指します。

■ 発表一覧

タイトル：抽象度が高いクエリによるアンケートデータの設問検索

著者：田中 稔也, 熊谷 雄介, 藤井 遼（株式会社博報堂DYホールディングス）

概要：膨大なアンケートデータから分析目的に合致する設問を検索する作業は、マーケターにとって手間と時間がかかる工程の一つです。本研究では、マーケターの曖昧な意図にもとづきアンケートデータの設問を検索する「設問検索」タスクを提案しました。マーケターのアノテーションによって構築した設問検索データセットに対する実験の結果、文埋め込みを用いた検索手法が高い精度を記録したものの、設問検索特有の課題も明らかになりました。

タイトル：Open Weight LLMs in Out-of-Distribution Setting: Search Ad Title Generation

著者：Tolmachev Arseny, Foran Joseph, 星野 智紀, 森川 裕介（株式会社博報堂テクノロジーズ）

概要：大規模言語モデル (LLM) は、一夜にして自然言語処理 (NLP) の領域に革命をもたらしました。しかし、依然として分布外 (OOD) シナリオにおいて課題を抱えています。本研究では、OOD シナリオの一例として、広告タイトル生成タスクにおける LLM の性能をケーススタディとして分析します。その結果、LLM は当初の予想を上回る性能を示しました。特に、命令調整 (instruction tuning) されたモデルは、非調整モデルと比較して顕著に安定していることが確認されました。一方で、蒸留された Llama 3.2 は、基盤モデルである Llama 3.1 よりも大幅に低い性能を示しました。さらに、汎用的な対話向けに命令調整されたモデルは、非調整モデルと比較して良し悪しの判断が難しい結果を示しました。

タイトル： Loss as a Data Introspection Method: Looking into Japanese Advertising Text Generation
著者： Joseph Foran, Arseny Tolmachev (Hakuhodo Technologies)

概要：本論文では、日本語広告文生成モデルにおける学習損失値の分布を検証したケーススタディを提示します。LongT5 アーキテクチャを用いて、高損失値および低損失値を示す学習例の特徴を分析しました。我々の調査により、以下の主要なパターンが明らかになりました：低損失の例では、反復的なフレーズや標準化された広告用語が頻出する一方、高損失の例では、より複雑な文法構造や自然言語パターンが特徴的に見られました。また、学習データの品質に関する潜在的な問題を特定し、モデルのパフォーマンスへの影響について考察を行いました。学習データにおける事例ごとの学習損失値の測定は、モデルの特性をより深く理解するための有用な診断ツールであることが判明しました。

タイトル：大規模言語モデルは翻訳品質を制御できない

著者：熊谷 雄介, 藤井 遼 (株式会社博報堂 D Y ホールディングス)

概要：機械翻訳の品質を解釈するには、「どの程度評価指標が低下したらどの程度翻訳品質が低下したのか」を定量的に議論しなければなりません。そのためには、低品質な訳を品質を制御しながら自動生成する必要があります。この研究では大規模言語モデル (LLM) による翻訳品質の制御に取り組みました。指示文に英語能力を指定した英日翻訳実験の結果、LLM の翻訳品質は定量的・主観的に低下しませんでした。また、「幼稚園児の英語力」などを指定した LLM は英文の内容を幼児語で答えました。これは、LLM が能力と文体を区別できていないことを示唆しています。

< 言語処理学会第 31 回年次大会 (NLP2025) 概要 >

日時：2025 年 3 月 10 日 (月) ~ 3 月 14 日 (金)

主催：一般社団法人 言語処理学会

会場：出島メッセ長崎 (〒850-0058 長崎県長崎市尾上町 4-1)

参加には、参加登録用ウェブサイトからの登録が必要です (有料)

NLP2025 公式サイト： <https://anlp.jp/nlp2025>

【本件に関するお問い合わせ】

株式会社博報堂 D Y ホールディングス

グループ広報・IR 室 倉品・西川・山崎

koho.mail@hakuhodo.co.jp

株式会社博報堂テクノロジーズ広報担当

hr-koho@hakuhodo-technologies.co.jp