

# ABCI 3.0開発加速利用（2025年度）成果概要（公開用）

課題名：大規模言語モデル等の透明性・信頼性の確保に向けた研究開発	実施時期：2025年11月～2026年3月 所属機関名：国立情報学研究所 代表者氏名：黒橋 禎夫
<p>成果概要：チューニングコンペティション（FT-LLM2026）を開催し、その中で希望したチームが計算資源を利用した。その結果、いずれのタスクにおいても計算資源を活用したチームが顕著な成果を示した。LLM-jpモデルによる中高数学タスクでは、上位チームが92.3%、84.7%と高精度を達成した。また、自由形タスクでは、計算資源を活用したチームが音声合成モデルを構築し、審査員評価で1位を獲得した。さらに、成果は言語処理学会第32回年次大会（NLP2026）のワークショップ（主催）で共有され、実践的知見の展開に寄与した。これらの結果から、計算資源の提供がモデル性能向上と応用開発の促進に有効であることが示された。</p>	
<p>成果のポイント：大規模言語モデルの透明性・信頼性の確保に向けた研究開発の一環として、国立情報学研究所が開発するLLM-jpモデルを対象にチューニングコンペティションを実施した。本コンペティションでは、数学能力向上を目的としたチューニングタスクと、応用事例の創出を目的とした自由形タスクを設定し、希望チームに計算資源を利用させることで実践的な開発環境を提供した。数学タスクでは透明性・再現性を重視し、オープン枠とクローズ枠を設け、オープン枠に限定して計算資源を利用させた。その結果、上位3チームはいずれも計算資源提供チームであり、精度92.3%、84.7%、82.9%と高水準を達成した。特に1位チームは、多様なデータ収集・精選やフィルタリング、モデル蒸留などの手法を組み合わせることで顕著な性能向上を実現した。さらに、これら上位チームのデータセットおよび再現手順は公開され、日本語オープンLLMの数学性能向上に資する知見が蓄積された。自由形タスクでは制約を設けず6チームに計算資源を提供し、発表内容に基づく審査により評価を行った。その結果、1位チーム（計算資源提供チーム）は音声埋め込みモデルの統合や既存データと合成音声を組み合わせた学習により、高い性能と独自性を示した。以上より、計算資源の提供が高性能モデル開発の促進に加え、再現性・透明性を備えた知見の蓄積および多様な応用創出に有効であることが示された。さらに、本コンペティションの成果を広く共有するため、言語処理学会第32回年次大会（NLP2026）においてNIIがワークショップを主催し、参加チームによる成果報告および手法の発表を実施した。これにより、チューニング手法やデータ構築・評価に関する実践的知見がコミュニティ内に展開され、日本語LLMの性能向上および透明性・再現性に関する取り組みのさらなる発展に寄与した。</p>	
<p>成果についてより詳細な情報を提供しているWebページ、発表論文などの情報： <a href="https://llm-jp.github.io/tuning-competition/index.html">https://llm-jp.github.io/tuning-competition/index.html</a></p>	