2025 年度年次報告書

先端国際共同研究推進事業
【2025 年度採択　日英共同公募　AI 分野】

【研究課題名】
安全な生成 AI のためのニューロシンボリック融合議論システム: NeSyDebates

研究代表者：
【佐藤　健】（情報・システム研究機構 人工知能法学研究センター センター長）

主たる共同研究者：
【Nguyen Le Minh】（北陸先端科学技術大学院大学（JAIST）教授）
【角田美穂子】（一橋大学　教授）

相手側研究代表者：
【Francesca Toni】（Imperial College Professor）

1. 研究進捗状況と成果

本研究の研究構想は、生成 AI におけるハルシネーションや規範違反の問題に対応するため、自然言語の政策・規制・法律（PRL）から機械可読な議論構造（CLA）を抽出し、新規事例に適用して違反の検出・分析・回避を行うニューロシンボリック融合議論システム（NeSyDebates）を構築する点にある。

研究成果としては、PRL から、法的知識表現言語 PROLEG 形式の LLM を用いた自動変換手法や事例からの事実抽出に基づく準拠チェック手法を開発するとともに、PROLEG と仮定ベース議論（ABA）の形式的対応関係を明らかにし、法的デバッグを「Contestable AI」へと拡張した。また、国際ネットワーク構築・拡大では、2026 年 1 月に Imperial College London でキックオフミーティングを実施し、日英の研究者間で研究構想を共有することで、各研究の相互関係を整理し、国際共同研究体制を強化した。若手研究者の人材育成としては、博士研究員を同大学に派遣し、PROLEG と ABA の橋渡し研究を推進した。その結果、ABA の争点性と PROLEG の法的デバッグの対応関係を示し、計算法学を動的に更新可能なシステムとして捉える新たな視点を提示するとともに、共同論文計画や国際的研究コミュニティ間の連携強化につながった。

2. 代表的な成果事例

1. Zhihan Cao, Fumihito Nishino, Hiroaki Yamada, Nguyen Ha Thanh, Yusuke Miyao, Ken Satoh. 2026. JBE-QA: Japanese Bar Exam QA Dataset for Assessing Legal Domain Knowledge. In Proceedings of the 2026 Language Resources and Evaluation Conference (LREC2026), to appear. Mallorca, Spain. European Language Resources Association.

概要：日本の司法試験短答式問題を XML 化し、それに基づいて LLM の性能を評価する手法を提案した。

2. Do Dinh Truong, Le Nguyen Khang, Nguyen Le Minh, UniSpec: Training-Free Speculative Decoding for Robust LLM Acceleration Across Languages and Hardware, to appear in The 64th Annual Meeting of the Association for Computational Linguistics (ACL 2026)

概要：speculative decoding に基づく非常に高速な手法を開発し、LLM の推論速度を向上させることができた。