

LLM生成テキスト検証サービスWISDOM-CE

～ LLMの出力とWeb上の多様な情報との整合・非整合のチェックによる「信頼できるAI」の実現 ～

概要

- LLMの生成テキストを膨大なWeb情報と照合し、妥当性を多面的に検証
- 生成テキストを支持する情報、矛盾する情報の両論提示で人の判断を支援
- LLMを動的に評価・強化する「能動的評価基盤」の中核技術として活用

WISDOM-CE: LLM生成テキスト検証サービス

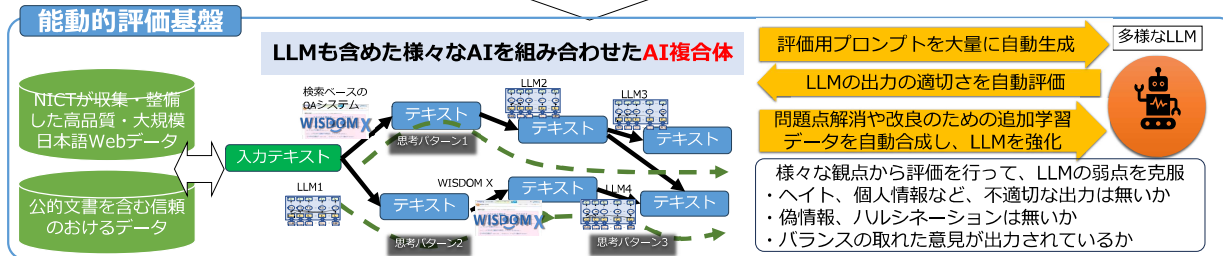
小中一貫校が増加した場合のメリットは、小中一貫校では、小学校と中学校の教員が協力して指導に当たることができるため、教育内容の充実や教員の指導力の向上につながることを期待できます。また、小中一貫校では、小学校と中学校の施設を一体的に整備することができるため、施設の有効活用や建設費の削減につながることを期待できます。さらに、小中一貫校では、小学校と中学校の教育課程を一体的に編成することができるため、教育内容の充実や教育活動の活性化が期待できます。

入力されたテキストに対して、

テキストの断片ごとに、支持等の整合情報(実線)、矛盾・反論等の非整合情報(点線)の存在を提示

下線クリックで関連情報の詳細を表示

関連情報をアドバイスとして挿入した文章も生成



特徴

- 単なる類似検索ではなく、支持・理由・帰結、矛盾・反論等の関係を持つ情報をWebから抽出して提示
- 「正しいか/誤りか」の二元論にとどまらず、複合的な視点を提供することで、人による意思決定を支援
- 20年に亘り収集したWeb736億ページから抽出した最大45.5TBの高品質大規模日本語データを活用

ユースケース

- LLMの出力に対して根拠・リスク等を提示することでハルシネーション等を特定、安心安全を担保
- 行政や報道など、出力テキストに対して特に透明性・説明責任が要求される分野でのLLMの業務利用に活用

今後の展開

- 関連情報の量・質、支持/反論の強さ等を定量化し、生成テキストごとの評価スコアを提示する手法を開発
- 多様なLLMを動的に評価し、弱点を自動的に改善する「能動的評価基盤」の中核技術として強化
- 民間企業等との共同研究を通じて社会実装、国産LLMの安全性担保や高機能化・高精度化にも貢献

