

## 優秀賞

D3 チームは、現在産業界で盛んに構築、試行されている RAG システムの評価フレームワークを構築しました。RAG においては、検索機構と GPT などの生成機構を組み合わせ、業務知識の問い合わせ応答などを実現する構造になっており、その構造を踏まえた RAGAS などのテストフレームワークが注目されています。

しかし、一つ一つの入力に対して多数の指標での評価が出力されても、ステークホルダーによる受け入れの判断や改善の検討につなげることができません。D3 チームはこの点を重要な課題ととらえ、ステークホルダーへの説明を目的とした評価フレームワークを構築しました。RAGAS に含まれない信頼度の観点を導入したほか、業務知識の種別による品質の違いなどの可視化、問題の原因分析も仕組みとして取り入れました。

タイムリーな課題にソフトウェアエンジニアリングらしい観点から取り組んだ点、提案した評価手法の評価もしっかりと行っている点が高く評価されました。

以上を高く評価して優秀賞を与えることとします。

国立情報学研究所 GRACE センター長・准教授

石川冬樹