

LLMを利用した電子実験ノートブック 自然言語処理による記録の構造化

日本ウォーターズ株式会社

安倍 雄一郎

yuichiro.a30@gmail.com

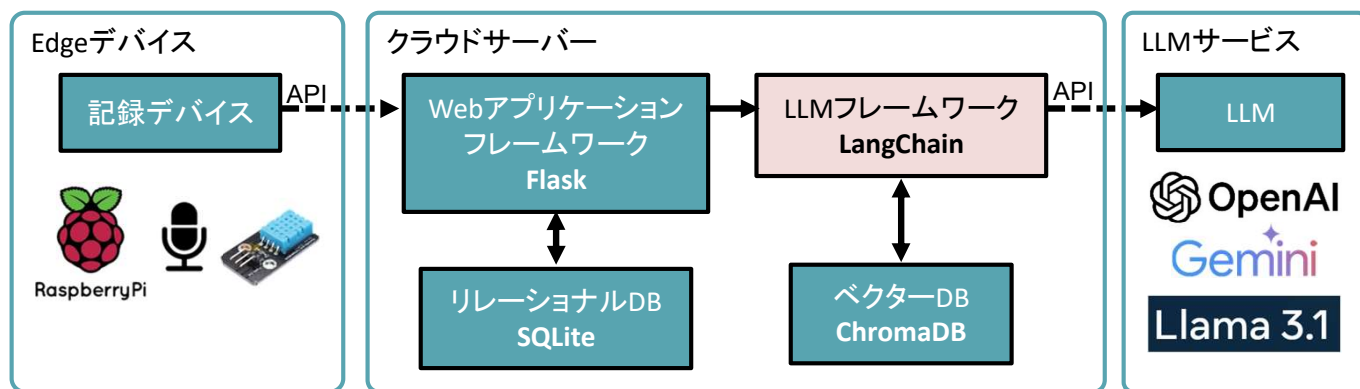
開発における問題点

- 近年、先行研究の実験結果が再現できない問題が深刻化しており、論文の撤回や、リソースの無駄な浪費に繋がっている。
- 研究者は実験の記録を正確に残すように教育されているが、実際にはプレッシャーから丁寧に記録を残すことができず、データの欠落、散逸、失敗データの消失を引き起こす。

手法・ツールの適用による解決

- 実験と同時並行で記録することで記録を残すハードルを下げる。
音声認識
- 音声記録を後で再利用できる形式に構造化する。
LLMによる自然言語処理

システム構成図



タスク

音声認識

- 文字起こし
- 認識精度向上

翻訳・要約

- 自然言語処理のため英語へ翻訳
- 実験記録から要約を作成
- プロンプトエンジニアリング
- 適切な変換事例の自動選択

推論による属性抽出

- 実験の再現性に重要な要素を設定して、LLMによる推論

構造化と可視化

- json構造へ型変換とバリデーション
- 結果を構造化された記録として表示

まとめ

