

## RAGによる既知障害情報の活用方法に関する考察

**研究コース5 人工知能とソフトウェア品質 AI4BUGDOCSチーム**

**研究員：手倉森 祥太（株式会社日立システムズ）**

**横内 雅人（株式会社日立ソリューションズ・クリエイト）**

**氏家 弘貴（キヤノンメディカルシステムズ株式会社）**

**狩野 薫（リコーITソリューションズ株式会社）**

**玉井 安明（株式会社デンソー）**

**主 査：石川 冬樹（国立情報学研究所）**

**副主査：徳本 晋（富士通株式会社）**

**アドバイザー：栗田 太郎（ソニー株式会社）**

# 背景 (1/3)

大量の障害情報から  
手動で探すのが大変。。

キーワード検索しても  
見つからないなあ。。

この障害って  
他のプロジェクトでも  
起きてない？



これらを今流行りのAI  
を使って解決できない  
かな？

目指す姿

A I を使った迅速な障害情報の検索

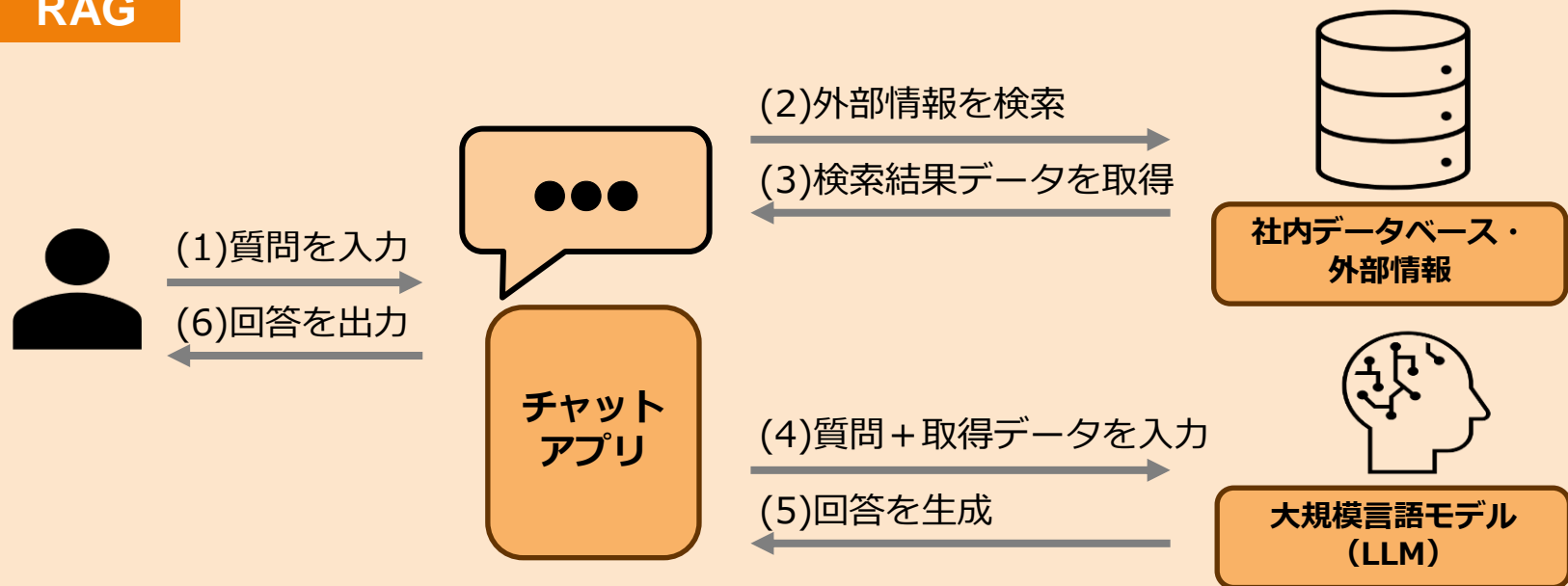
# 背景 (2/3)

目指す姿に近づくアプローチとして、生成AIの応用技術のひとつ【RAG】に着目



RAG(Retrieval-Augmented Generation)とは、外部データベースから情報を検索し、その情報と質問内容を基に生成AI(LLM)に回答させる手法のこと

## RAG



# 背景 (3/3)

RAGを用いて障害情報を検索することにより、  
以下の面で効果が期待できる

## 1. 効率面

手動検索の代わりに、生成AIから必要な情報を自動的に取得できる

## 2. 品質面

生成AIにより文脈を考慮した関連性の高い情報を取得できる

# 研究テーマ (1/3)

## LLMが答えられる質問

チャレンジャー号爆発事故の原因について教えてください。

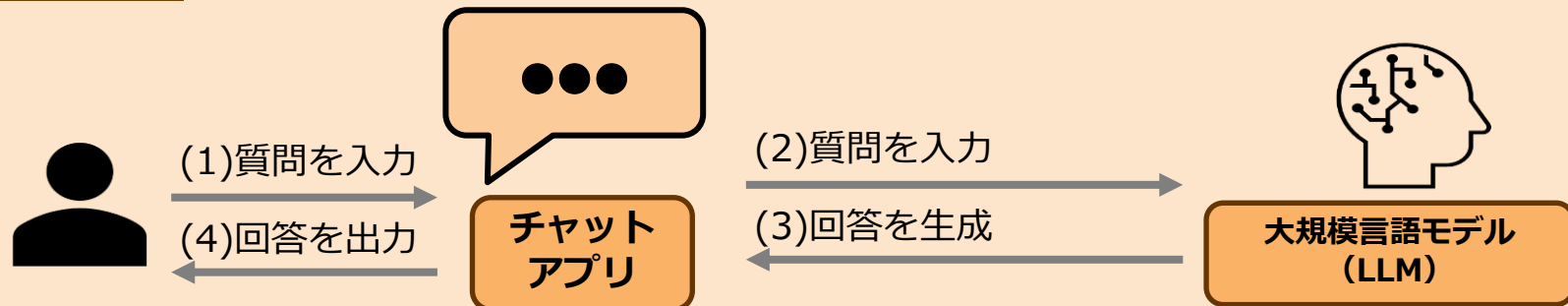
チャレンジャー号爆発事故の原因は、XXXXXXXX  
XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX  
XXXXXXXX XXXXXXXX XXXXXXXX XXXXXXXX

## LLMが答えられない質問

うちの会社内での障害事故情報一覧を出してください。

そんな情報は持ってないから分からないよー！！！！

## LLM



**期待した回答が返ってこない**

# 研究テーマ (2/3)

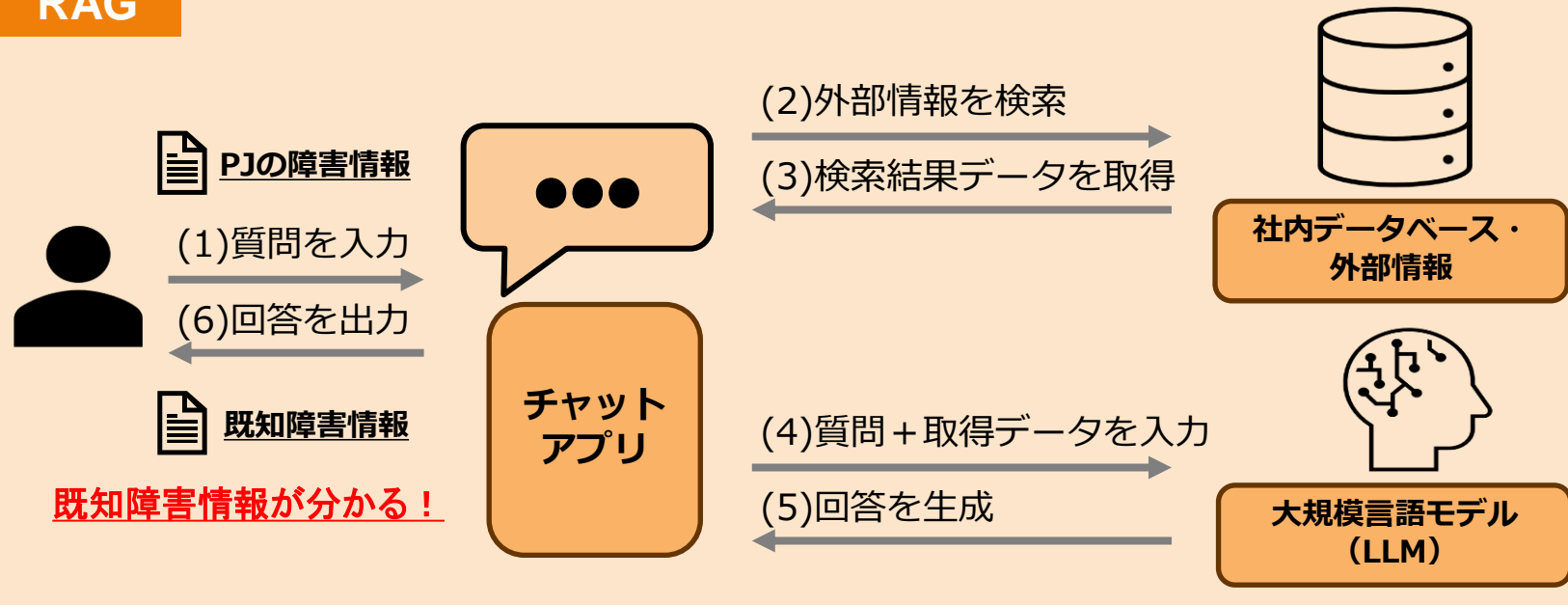
## RAGが答えられる質問

うちの会社内での障害事故情報一覧を出してください。

おっけー!!!!!!  
会社の障害一覧を出力するねー!!!

わーい!!!うれしい!!!

## RAG

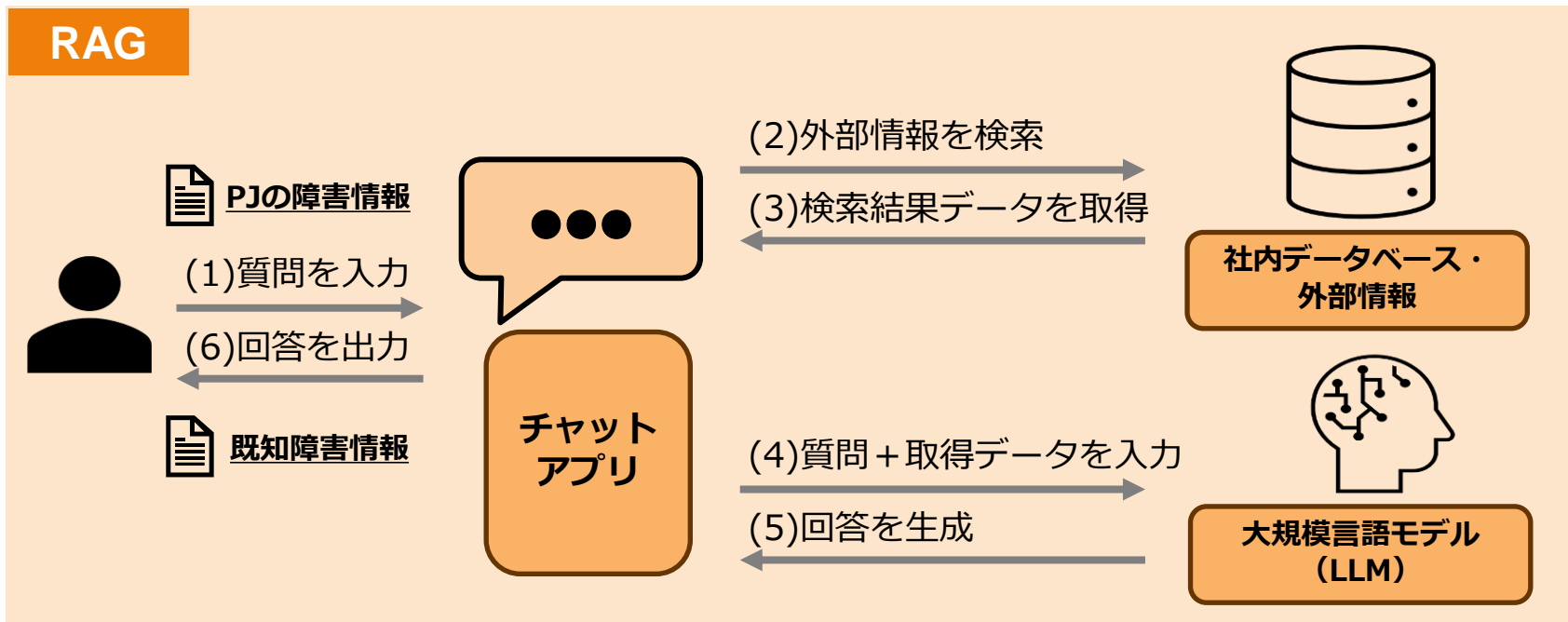


# 研究テーマ (3/3)

RAGを採用し、実際に既知障害情報の検索を行った場合、どのような条件でどの程度の効果が得られるかが不明確

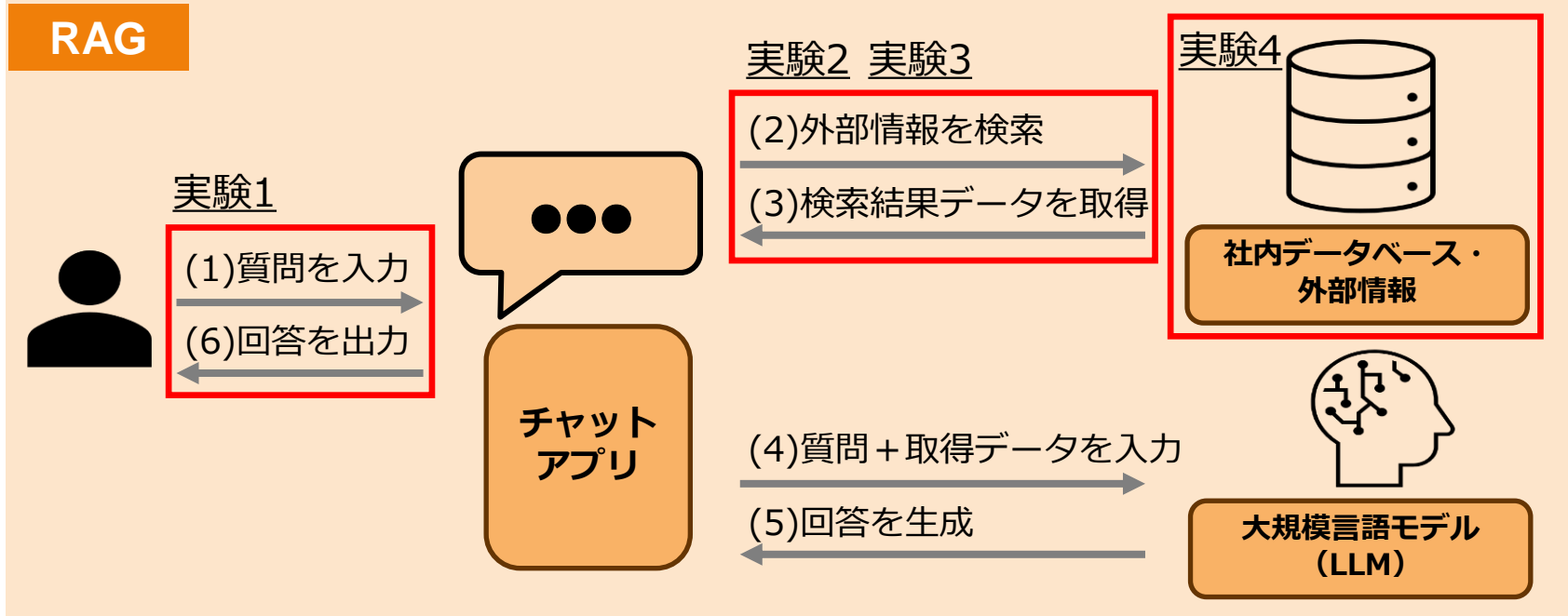


**RAGの検索性能と回答精度を評価し、向上させる方法を探る！**



# 研究方法

No	実験内容
1	LLMに障害情報を与えると回答にどう影響するか
2	データの形式の違いが、RAGの検索にどう影響するか
3	同一障害に関する質問の違いが、RAGの検索にどう影響するか
4	データの形式や構造の違いが、検索性能と回答精度にどう影響するか





# 実験1 (1/2)

概要: LLMに障害情報を与えると回答にどう影響するか

方法: ①障害情報ありと②障害情報なしで、LLMの回答を比較

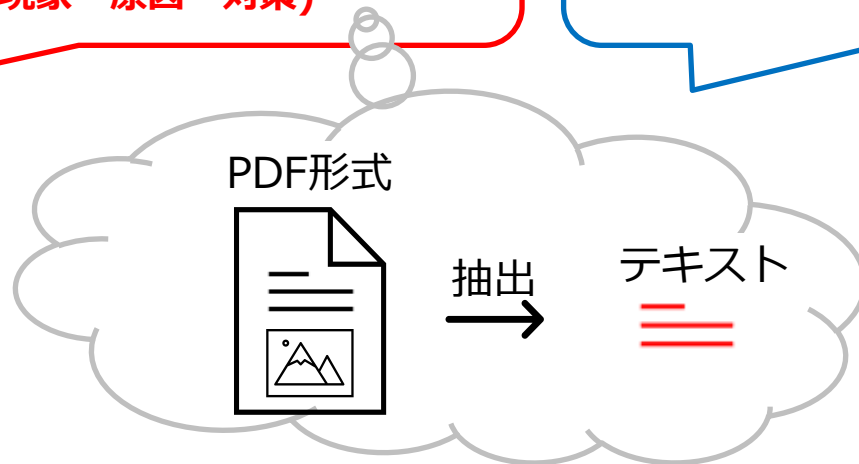
- LLM: ChatGPT-4o
- データ: IPAが公開する障害情報 (現象+原因+対策) (53件)
- プロンプト (2パターン)

## ①障害情報あり

以下の事例をもとに、  
(現象)の原因を教えてください。  
その原因に対する対策を教えてください。  
■ 障害情報(現象・原因・対策)

## ②障害情報なし

(現象)の原因を教えてください。  
その原因に対する対策を教えてください。



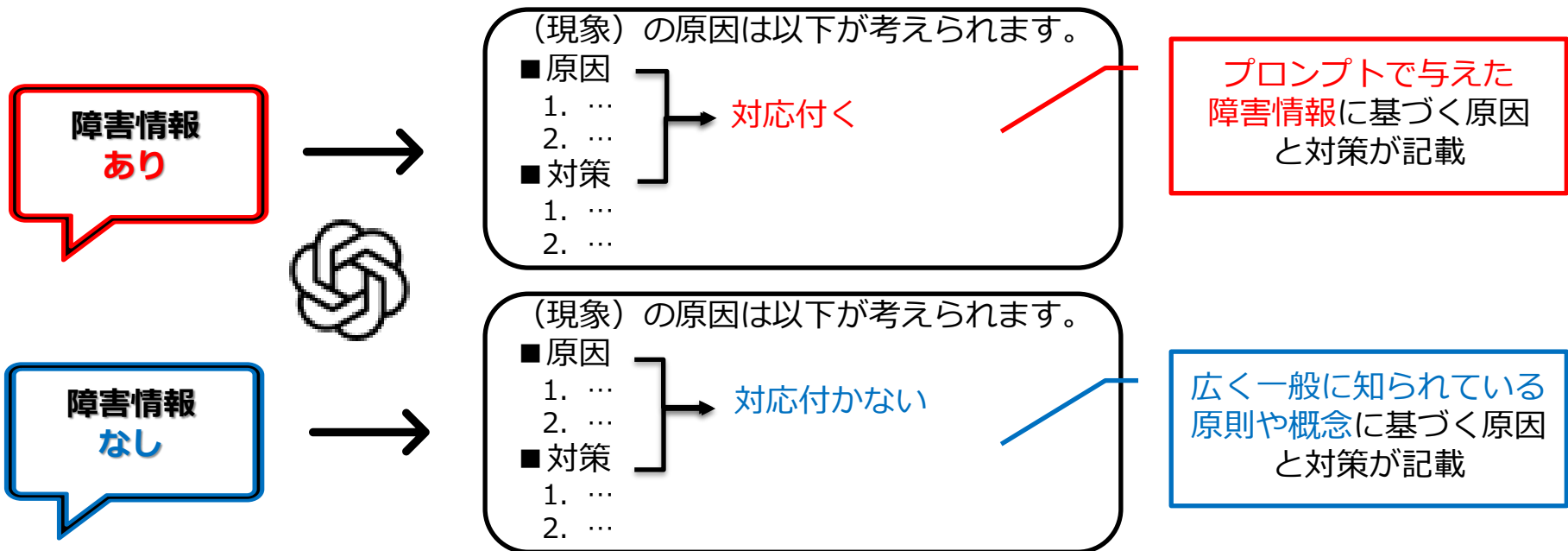
# 実験1 (2/2)

概要: LLMに障害情報を与えると回答にどう影響するか

結果: 障害情報なしでも一定の回答が可能だが、内容に違いあり

## ①整合性の違い

## ②根拠の違い



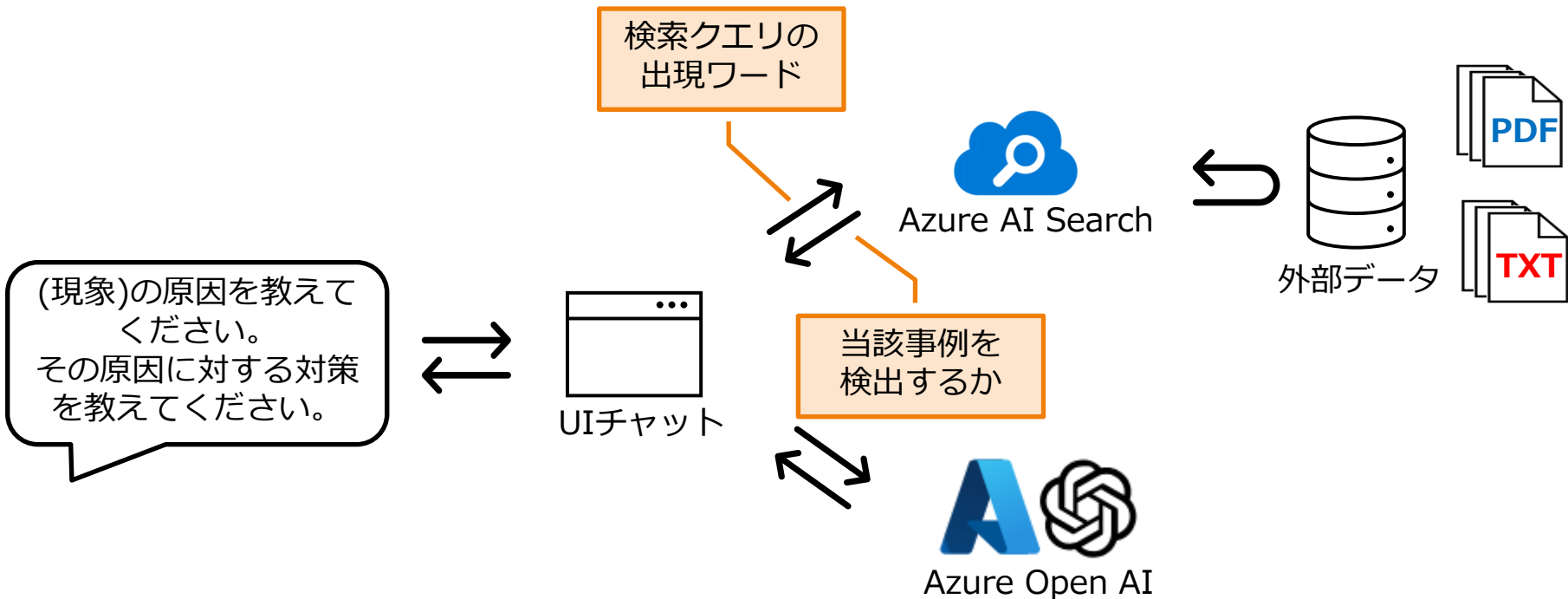
障害対応に役立つ原因や対策を得るには  
LLMに具体的な障害情報を提供すること

# 実験2 (1/2)

**概要:**データの形式の違いが、RAGの検索にどう影響するか

**方法:**2種のデータ形式を用い、生成する検索クエリ構成を比較

- プロンプト (ある1事例に基づく質問)
- Azure AI Search (セマンティック検索) + Azure Open AI
- 実験1で使用した障害情報 (各53件の①PDF形式、②TXT形式)

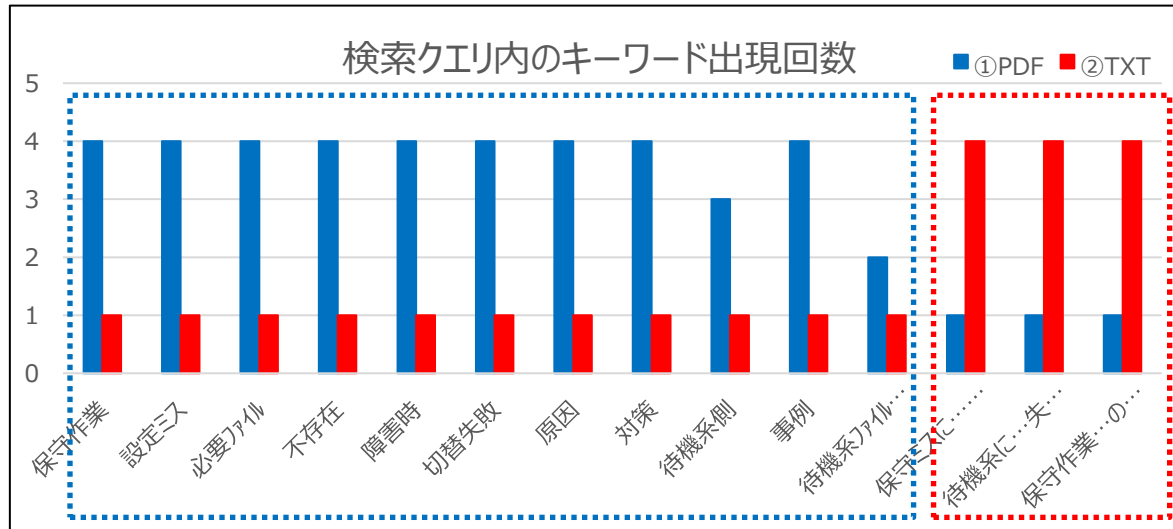


# 実験2 (2/2)

**概要:データの形式の違いが、RAGの検索にどう影響するか**  
**結果:検索クエリのキーワードに違いが生まれ、検索結果にも影響した**

①キーワードの表現の違い

N=5



②検出の違い

①短い単語は意味が曖昧になり事例の検出率が下がる  
 ②長いフレーズでは細かい表現の違いで検出漏れが発生しやすい

障害情報に関連するが、広く一般的に使われる単語や短いフレーズが多い

文章の一部をそのまま用いたフレーズが多い



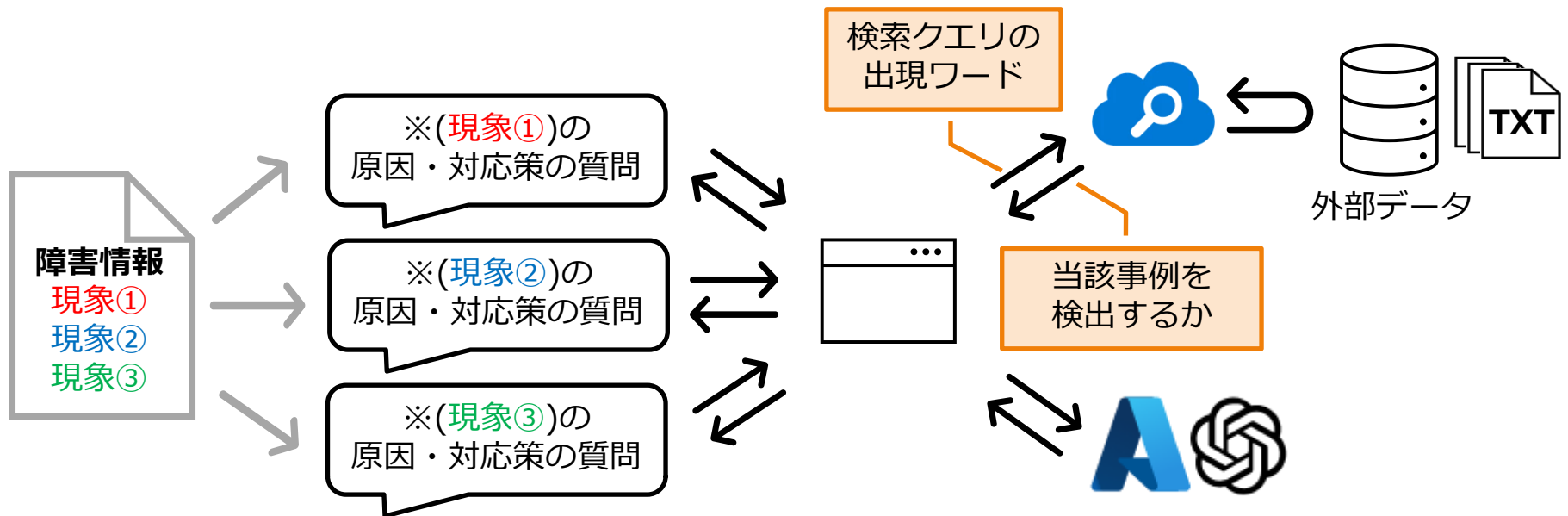
**障害情報のデータの形式の違いが検索クエリの質(フレーズの長さや意味のバランス)に影響する**

# 実験3 (1/2)

**概要: 同一障害に関する質問の違いが、RAGの検索にどう影響するか**

**方法: ある障害について3パターンの質問を行い、生成する検索クエリ構成を比較**

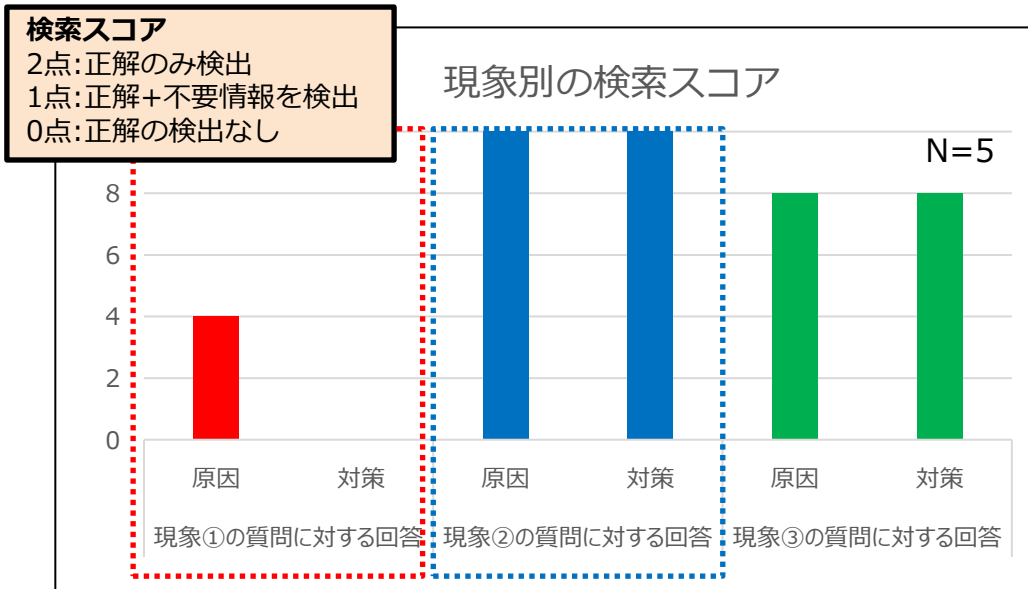
- 実験1で使用した障害情報 (53件のTXT)
- Azure AI Search (セマンティック検索) + Azure Open AI
- プロンプト (同一障害に関する質問 3パターン)



# 実験3 (2/2)

**概要:同一障害に関する質問の違いが、RAGの検索に  
どう影響するか**

**結果:各質問の検索クエリの出現ワードの希少性が、  
検索結果に影響した**



現象①の検索クエリの出現ワード

保守作業, 設定ミス, 必要ファイル, 不存在, 障害時, 切替失敗, 原因, 対策, 待機系側, 事例, 待機系ファイル不足, 保守ミスに…の原因と対策, 待機系に…事例, 保守作業…と対策

他の事例でも含まれるワードが多数

現象②の検索クエリの出現ワード

デスクモジュール, 自己診断機能, 異常検知, 機能停止, 事例, 原因, 対策, 仕様

1事例のみに含まれるワード

**障害情報とプロンプトに“特徴的なキーワード”を含めると  
検索精度が向上する**

# 実験4 (1/2)

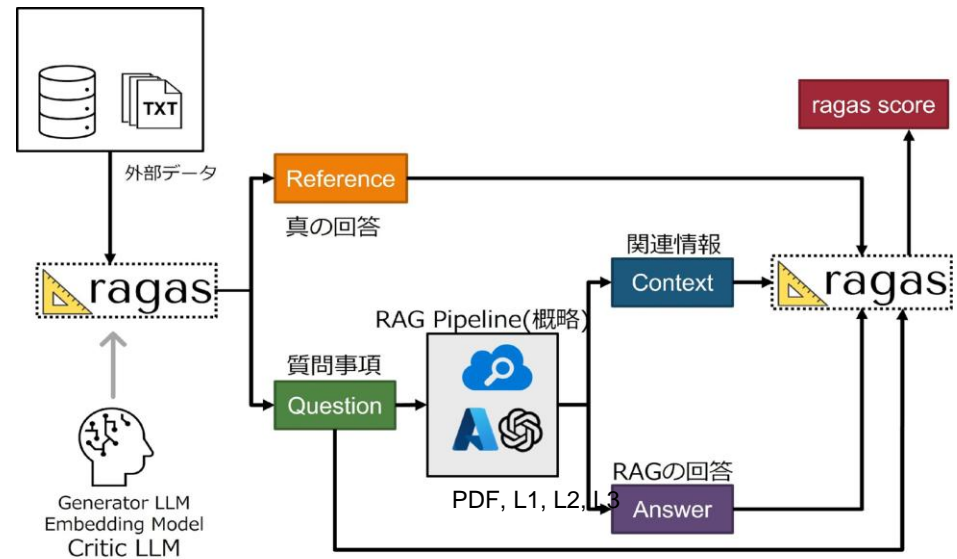
**概要:データの形式や構造の違いが、検索性能や回答精度にどう影響するか**

**方法:PDF、L1、L2、L3の4種類のデータ形式と構造を使い、RAGによる出力結果の精度を比較した**  
**RAGAS**のテストセット生成機能を使い、ランダムな質問83件に対するレスポンスを定量評価した

RAGAS(Retrieval Augmented Generation Assessment)は、RAGを評価するツールを提供するライブラリ  
 ⇒検索性能と回答精度を評価可能

## ※データ群の詳細

- PDF(非構造化データ(図表含む))
- L1(PDFのテキスト化データ(図表除く))
- L2(L1を整形したテキストデータ)
- L3(L2をタグ付けした構造化データ)

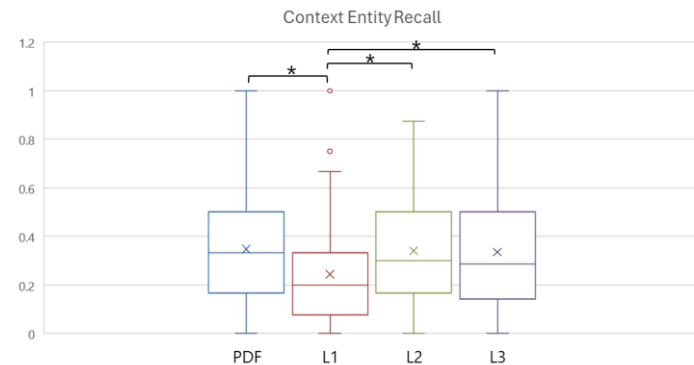
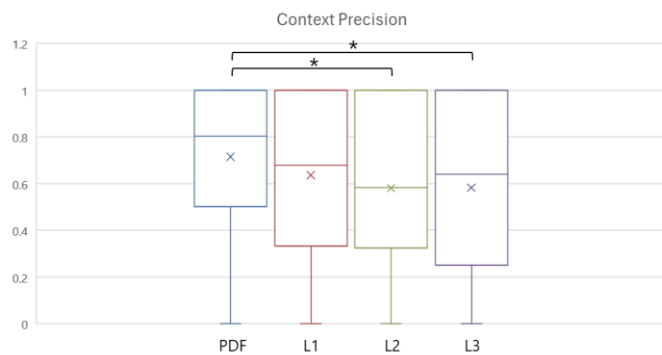


# 実験4 (2/2)

**概要:データの形式や構造の違いが、検索性能や回答精度にどう影響するか**

**結果:データ形式や構造は検索性能に影響を与えるが、回答精度には大きな影響を与えなかった**

- コンテキストの精度 (**Context Precision**): PDFはL2、L3よりも有意に高い
- 回答に必要な単語・キーワード網羅性(**Context Entity Recall**): L1が低い



RAGASで、検索性能と回答精度を定量的に評価できる

検索性能ではPDF形式がTXT形式より優れていた

回答精度の低下は**セマンティック検索**や推論能力の高いLLM利用で抑制できる



# 考察

## 考察:

**RAGの採用で障害に対する回答精度の向上が期待できる**  
→ 検索性能に大きな影響を与えるデータ形式や構造を考慮  
することが必要

## 現場活用での留意点:

**PDF形式は有利であり、運用には検索精度を確保しつつ、  
回答性能を維持する仕組みが効果的**

回答精度を高めるためには、  
障害情報記載時に特徴的なキーワードの盛り込みが重要

- ・ドキュメントの構造
  - ・文節区切りを適切に設計する
- ⇒ 効率的に検索精度を向上させ、回答の質を保てる

# 今後の展望

- **検索性能や回答精度を向上する障害情報の構造の検討**
  - 研究では社内に一般的に存在するデータ形式を想定し検討したが、より適切なデータ形式、構造化の選択によりRAG性能を向上する
- **最適なデータ整理方法の検討**
  - 障害情報における「問題・原因・対策」といった構成が検索性能や回答精度にどのような寄与をするか評価し、情報取得しやすいデータ蓄積/変換をする必要がある
- **マルチモーダルRAG(音声、画像、動画等)の検討**
  - 多様な障害情報の表現に対する対応による回答精度向上

# 謝辞

**本研究にご協力頂いた全ての皆さまに感謝いたします。**

特に石川冬樹主査、栗田太郎副主査、徳本晋副主査には  
研究に際し、多大なるご支援、的確なご指導を賜りました。  
深く御礼を申し上げます。

**ご清聴ありがとうございました。**