

令和7年度  
中小企業庁 取引調査員活動支援AI事業  
調査報告書

2026/2/27



Build Beyond As One.

# 目次

## 1. 本事業概要および目的

---

- 1.1. 実施目的および概要
- 1.2. エグゼクティブサマリ

## 2. 本事業のアプローチ

---

- 2.1. 作業全体スケジュール
- 2.2. 成果物

## 3. 入力補助機能の有効性検証

---

- 3.1. 検証の実施方針
- 3.2. ヒアリング結果記録アプリの全体像
- 3.3. As-Is/To-Beの概要
- 3.4. ユースケースの概要
- 3.5. 各ユースケースの対象ヒアリング項目
- 3.6. 検証における評価観点と評価方法
- 3.7. 評価結果

## 4. 入力補助機能の環境構築および検証

---

- 4.1. 検証の実施方針
- 4.2. 検証環境の要件
- 4.3. 検証環境の構成・仕様
- 4.4. 基本機能と利用イメージ
- 4.5. 検証方法
- 4.6. 検証結果
- 4.7. 今後対応すべき課題

## 5. 業種別分析機能の環境構築および検証

---

- 5.1. 検証の実施方針
- 5.2. 検証対象ユースケースの概要
- 5.3. 検証環境の要件
- 5.4. 検証環境の構成・仕様
- 5.5. 基本機能と利用イメージ
- 5.6. 検証方法
- 5.7. 検証結果
- 5.8. 今後対応すべき課題

## 6. 今後の取り組み

---

- 6.1. 今後想定される課題と取り組み
- 6.2. 取引調査員におけるヒアリング業務のロードマップ

# 1. 本事業概要および目的

1.1. 実施目的および概要

1.2. エグゼクティブサマリ

## 1. 本事業概要および政策目的

### 1.1. 実施目的および概要

- 本事業の目的は、調査員の業務効率化・高度化のための最適な生成AIの業務への活用領域および適応方法を特定・確認することである。
- その目的実現のために、「入力補助機能の有効性検証」「入力補助機能の環境構築および検証」「業種別分析機能の環境構築および検証」を実施した。

#### 実施目的

調査員の業務を効率化・高度化するための、最適な生成AI（本事業ではテキスト生成を中心とした検証のため、以降はLLMで統一する）の業務への活用領域および適応方法を特定・確認すること

#### 実施概要



#### 入力補助機能の有効性検証

##### 概要

調査員のヒアリング記録業務を効率化・高度化するために、LLMを用いた入力補助機能のユースケースを整理する。その有効性を定性・定量評価の両面から検証する。

##### 主な実施内容

- ・ ヒアリング項目、課題、ニーズの整理
- ・ As-Is/To-Beやユースケースの作成・整理
- ・ 有効性検証におけるKPI設計
- ・ 有効性の定性・定量評価



#### 入力補助機能の環境構築および検証

##### 概要

ヒアリング結果記録アプリ（kintone）とのシステム連携を検証しつつ、LLM環境を実装する。調査員や職員の視点から入力補助機能が業務に活用できるか、また業務効率化や品質向上につながるかを検証する。

##### 主な実施内容

- ・ kintone-Azure間の連携検証
- ・ 検証用の環境開発
- ・ 検証方法の検討
- ・ 検証および検証結果の整理と課題抽出



#### 業種別分析機能の環境構築および検証

##### 概要

調査員の業種別分析業務を効率化・高度化するために、LLMを用いた業種別分析機能のユースケースを整理する。調査員や職員の視点から業種別機能が業務に活用できるか、また業務効率化や品質向上につながるかを検証する。

##### 主な実施内容

- ・ 業種別分析の課題、ニーズの整理
- ・ 業種別分析のユースケースの作成・整理
- ・ 検証用の環境開発
- ・ 検証方法の検討
- ・ 検証および検証結果の整理と課題抽出

# 1. 本事業概要および政策目的

## 1.2. エグゼクティブサマリ



### 入力補助機能の有効性検証

第3章  
第4章  
第5章

- 業務課題とニーズを整理
- As-Is/To-Beからユースケースを整理
- ユースケースの機能を定量・定性の両面から評価し、業務への有効性を検証の両面から検証する。

#### ユースケースの整理

調査員が入力したヒアリング記録に対するユースケースを整理

##### A. フリー入力文章の標準化

ヒアリング記録のフリー入力文章を標準化した結果を提示

##### B. 法令資料を参照しないカテゴリ値の入力・チェック効率化

ヒアリング記録情報を基にカテゴリ値の推論結果を提示

##### C. 法令資料を参照するカテゴリ値の入力・チェック効率化

ヒアリング記録情報と法令資料を基に法令に関連するカテゴリ値の推論結果を提示

##### D. フリー入力文章の整合性確認効率化

過去ヒアリング記録情報を基にフリー入力文章と過去情報の整合性を確認した結果を提示

#### 有効性の評価

業務での有効性が見込める精度とスコア

検証結果を基に機能要件を再整理し、環境構築へ移行

	定量評価		定性評価
	適合率	再現率	スコアリング
A	91%	93%	9.7/10
B	88%	88%	9.3/10
C	91%	93%	9.8/10
D	80%	80%	9.3/10



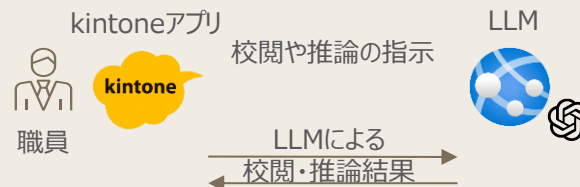
### 入力補助機能の環境構築および検証

第3章  
第4章  
第5章

- ヒアリング結果記録アプリ（kintoneアプリ）とのシステム連携を検証し、検証環境を実装
- 調査員と職員の視点から業務への適用可能性および導入効果进行评估

#### 検証環境の構築

kintoneアプリとのシステム連携を検証し、検証環境を構築



#### 検証結果および課題の抽出

- 評価された点
  - チェック作業効率化への有効性
  - 客観情報としての有用性
- 課題
  - 業務活用水準を満たす出力品質
  - 専門領域に適応した推論
  - 円滑な利用を実現するUX設計
  - LLMの適切活用の教育・運用

- ✓ 過半数の検証者が継続利用を希望
- ✓ 出力品質やUX改善など課題に対する対策案を検討し、次フェーズで実施する想定 (第6章に詳細記載)



### 業種別分析機能の環境構築および検証

第3章  
第4章  
第5章

- 分析業務におけるユースケースを整理
- 選定した業務における分析軸等を確認し、検証環境の構築。業務高度化に向けた適用可能性を検証

#### ユースケースの整理

蓄積されたヒアリング記録データに対するユースケースを整理

- 分析軸に沿ったヒアリング結果の抽出・要約・示唆提示  
指示や分析軸に沿ったヒアリング記録データの抽出・要約・示唆出し
- 条件一致ヒアリング結果抽出  
任意の抽出条件に合致するヒアリング記録データの抽出

#### 検証環境の構築

整理した機能要件に基づき、検証用アプリを構築（なお検証を円滑に進めるためにプロンプト作成部分はマクロで代用）



プロンプト作成マクロ



業種別分析機能アプリ

#### 検証結果および課題の抽出

- 評価された点
  - 過半数が継続利用を希望
  - 抽出や要約の出力は有用
- 課題
  - 柔軟性のある抽出
  - 出力結果の精緻化やUX改善
  - 既存システムとの差別化

## 2. 本事業のアプローチ

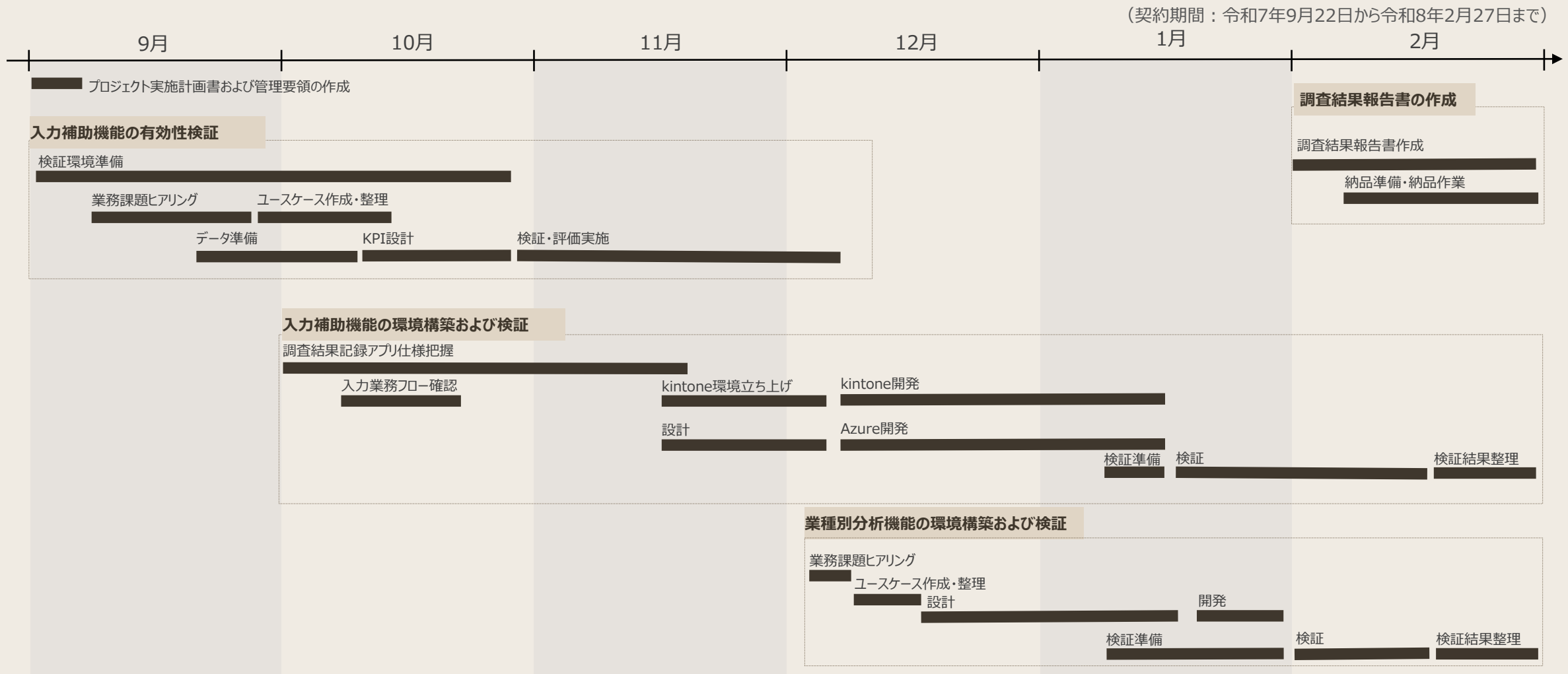
2.1. 作業全体スケジュール

2.2. 成果物

## 2. 本事業のアプローチ

### 2.1. 作業全体スケジュール

- 入力補助機能の業務適用の有効性検証を実施し、入力補助機能の検証用環境構築および検証を行った。
- さらに業種別分析機能の業務適用方法の検討を実施し、業種別分析機能の検証用環境構築および検証を行った。



## 2. 本事業のアプローチ

### 2.2. 成果物

- 以下が本事業での成果物である。

成果物分類	成果物名	ファイル名
調査報告書等一式(非公開用)	調査報告書	01_01_調査報告書.pptx、01_01_調査報告書.pdf
	報告書骨子(様式1)	01_02_報告書骨子(様式1).docx
	委託調査報告書公表用書誌情報(様式2)	01_03_委託調査報告書公表用書誌情報(様式2).xlsm
	二次利用未承諾リスト(様式3)	01_04_二次利用未承諾リスト(様式3).xlsx
	調査で得られた元データ	01_05有効性検証結果.xlsx、01_05入力補助機能・業種別分析機能アンケート結果.xlsx
調査報告書等一式(公表用)	調査報告書および二次利用未承諾リスト(公表用)	02_01_調査報告書.pdf
補足資料	業務実施計画書および実施要領	03_01_業務実施計画書および実施要領.pptx
	定例会議等の議事録	03_02_議事録.zip
	標準コーディング規約	03_03_標準コーディング規約.pptx
	設計書	03_04_システム環境構成図.pptx、03_04_検証環境パラメータシート.xlsx
	プログラム一式	03_05_プログラム一式.zip
	運用設計書・保守運用作業計画書等	03_06_保守運用作業計画書.pptx
	操作手順書(管理者向け)	03_07_操作手順書.pptx
	情報セキュリティ管理計画書	03_08_情報セキュリティ管理計画書.pptx
	引継ぎ資料	03_09_設計書_実装項目一覧.xlsx、03_09_課題一覧.xlsx

# 3. 入力補助機能の有効性検証

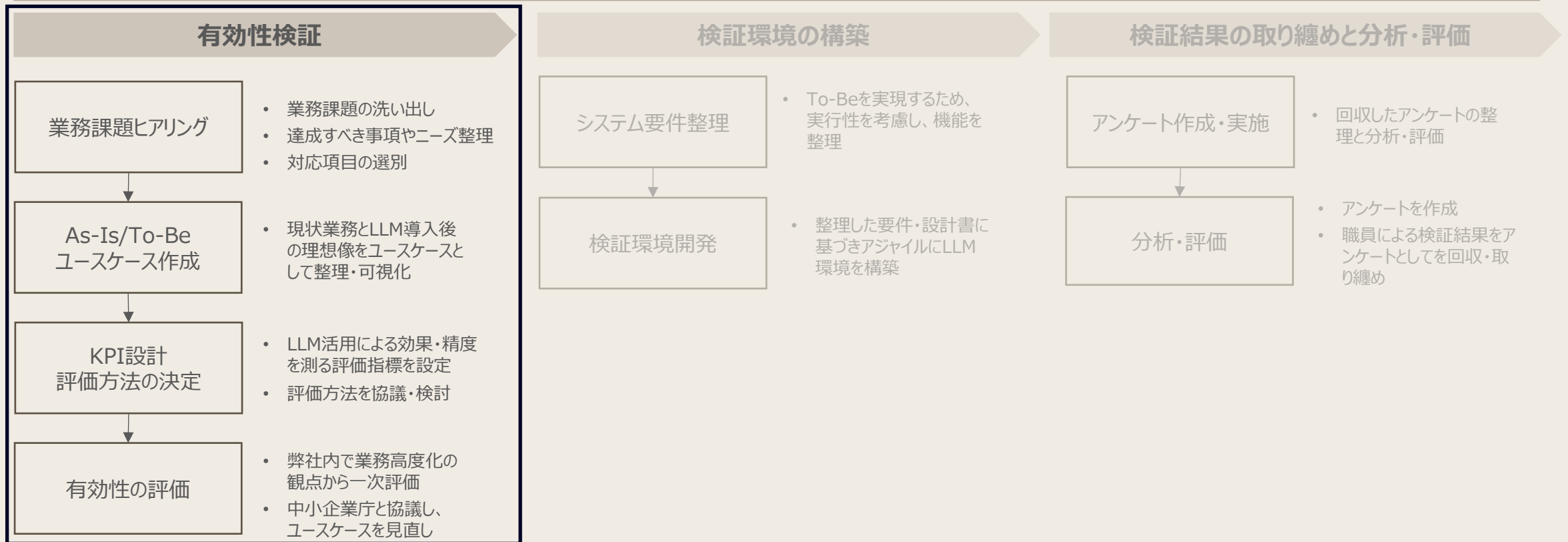
- 3.1. 検証の実施方針
- 3.2. ヒアリング結果記録アプリの全体像
- 3.3. As-Is/To-Beの概要
- 3.4. ユースケースの概要
- 3.5. 各ユースケースの対象ヒアリング項目
- 3.6. 検証における評価観点と評価方法
- 3.7. 評価結果

### 3. 入力補助機能の有効性検証

#### 3.1. 検証の実施方針

- 業務ヒアリングを通じて、実施内容の検討と優先度の整理を行い、将来像（To-Be）およびユースケースを定義した。弊社にて業務高度化に係る一次評価を行い、ユースケースの精緻化を実施した。

#### 入力補助機能の検証プロセス



### 3. 入力補助機能の有効性検証

## 3.2. ヒアリング結果記録アプリの全体像

- ヒアリング結果記録アプリは、大きくヒアリング先基本情報アプリ、取引先情報アプリ、テーマ別ヒアリング情報アプリから構成される。

kintone

### ヒアリング結果記録アプリ（kintone）



#### ヒアリング先基本情報アプリ

データ管理情報  
ヒアリング管理情報  
ヒアリング先中小企業基本情報  
自社が委託事業者となる場合の取引の状況  
要望事項  
備考  
関係制度の周知  
臨時調査項目



#### 取引先情報アプリ

取引先基本情報  
価格交渉等の状況  
備考



#### テーマ別ヒアリング情報アプリ

支払条件に関する情報  
価格決定方法に関する情報  
知財の保護について  
型取引について  
価格交渉フォローアップ調査  
ヒアリング結果に対する所見等

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 As-Is/To-Beの一覧

- ヒアリング項目の内容から、入力補助機能で実現可能と考えられるAs-Is/To-Beを5つ設定した。

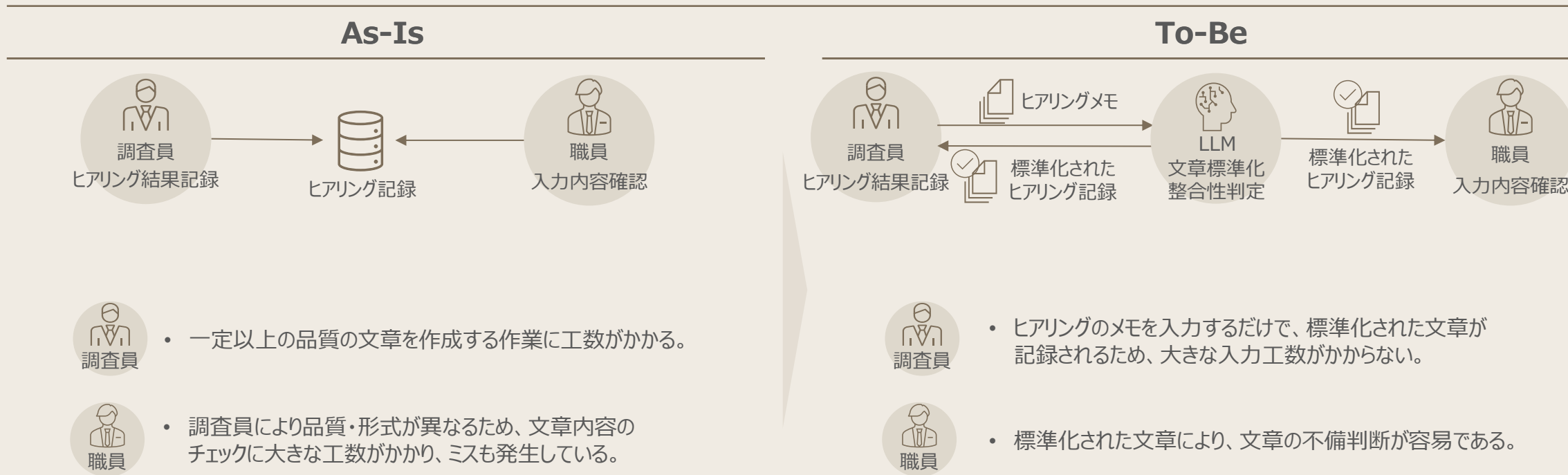
As-Is/To-Be名	As-Is概要	To-Be概要
① 文章入力とチェックの効率化	<ul style="list-style-type: none"><li>・ 一定以上の品質の文章作成に工数がかかる。</li><li>・ 入力文章が標準化されておらず、チェックに工数がかかる。</li></ul>	<ul style="list-style-type: none"><li>・ 情報を入力するだけで、標準化された文章が記録される。</li><li>・ 標準化された文章により、文章の不備の判断が容易である。</li></ul>
② 企業関連情報の 整合性チェックの効率化	<ul style="list-style-type: none"><li>・ 企業関連情報が正しいかの参照・確認に工数がかかる。</li></ul>	<ul style="list-style-type: none"><li>・ 企業関連情報に誤りがある可能性がわかり、確認が容易である。</li></ul>
③ 業種カテゴリの 入力とチェックの効率化	<ul style="list-style-type: none"><li>・ 業種の種類や定義が複雑で、業種の判断、チェックに工数がかかる。</li></ul>	<ul style="list-style-type: none"><li>・ 業種を判断するための情報が整理されていて、業種の判断、チェックが容易である。</li></ul>
④ 法令資料を参照する カテゴリ入力とチェックの効率化	<ul style="list-style-type: none"><li>・ 法令の参照範囲が広く、正確なカテゴリ値の判断、チェックに工数がかかる。</li></ul>	<ul style="list-style-type: none"><li>・ カテゴリ値を判断するための情報が整理されていて、カテゴリの判断、チェックが容易である。</li></ul>
⑤ 法令資料を参照しない カテゴリ入力とチェックの効率化	<ul style="list-style-type: none"><li>・ 他のヒアリング項目を参照する必要があり、カテゴリ値の判断、チェックに工数がかかる。</li></ul>	<ul style="list-style-type: none"><li>・ カテゴリ値を判断するための情報が整理されていて、カテゴリの判断、チェックが容易である。</li></ul>

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 ①文章入力とチェックの効率化

- ヒアリングのメモを入力するだけで標準化された文章が作成されることで、入力・チェックの工数が削減される。
- 企業情報の誤りの可能性や詳細がわかることで、チェックの工数が削減され、品質も向上する。

#### As-Is/To-Be案



#### 対象ヒアリング項目

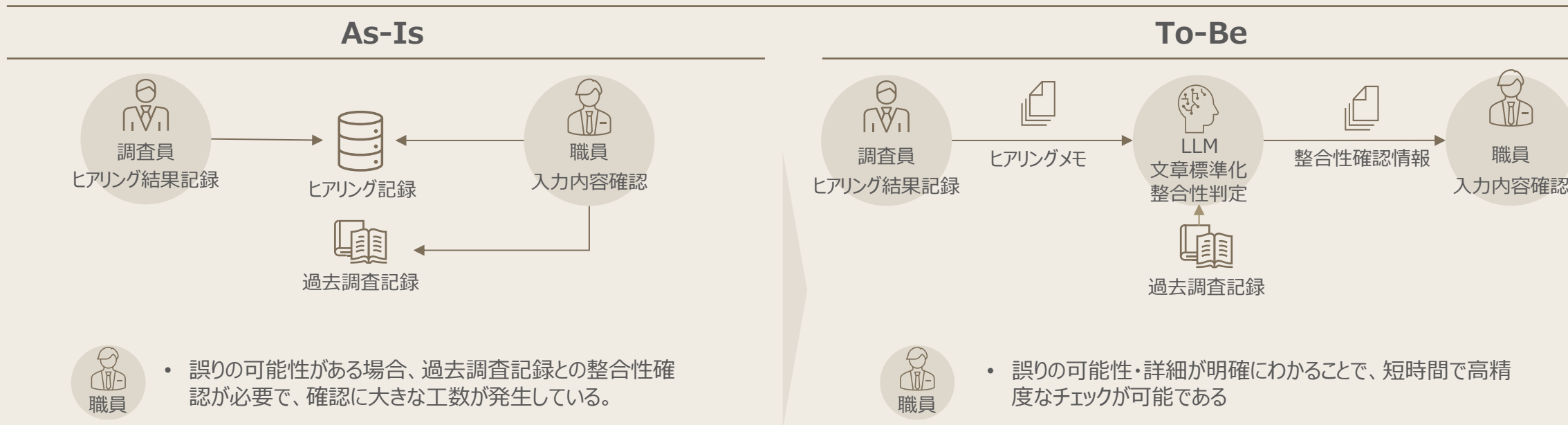
G	ヒアリング先基本情報アプリ	取引先情報アプリ	テーマ別ヒアリング情報アプリ
	ヒアリング先中小企業基本情報 自社が委託事業者となる場合の取引の状況	取引先基本情報	価格交渉フォローアップ調査 ヒアリング結果に対する所見等
	備考 臨時調査項目		

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 ②企業関連情報の整合性チェックの効率化

- 企業関連情報の誤りの可能性や詳細がわかることで、チェックの工数が削減され、品質も向上する。

#### As-Is/To-Be案



#### 対象ヒアリング項目

**G** ヒアリング先基本情報アプリ  
ヒアリング先中小企業基本情報

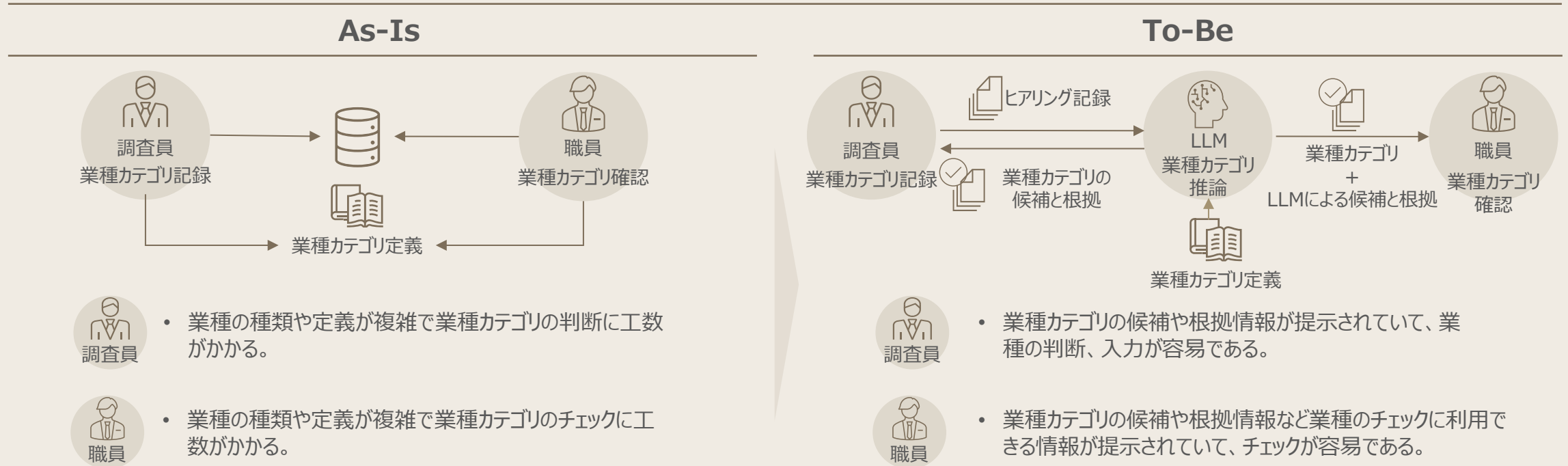
**取引先情報アプリ**  
取引先基本情報

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 ③業種カテゴリの入力とチェックの効率化

- 業種カテゴリの候補やその根拠が提示されることで、入力・チェックの工数が削減される。

#### As-Is/To-Be案



#### 対象ヒアリング項目

**G** ヒアリング先基本情報アプリ

ヒアリング先中小企業基本情報

**取引先情報アプリ**

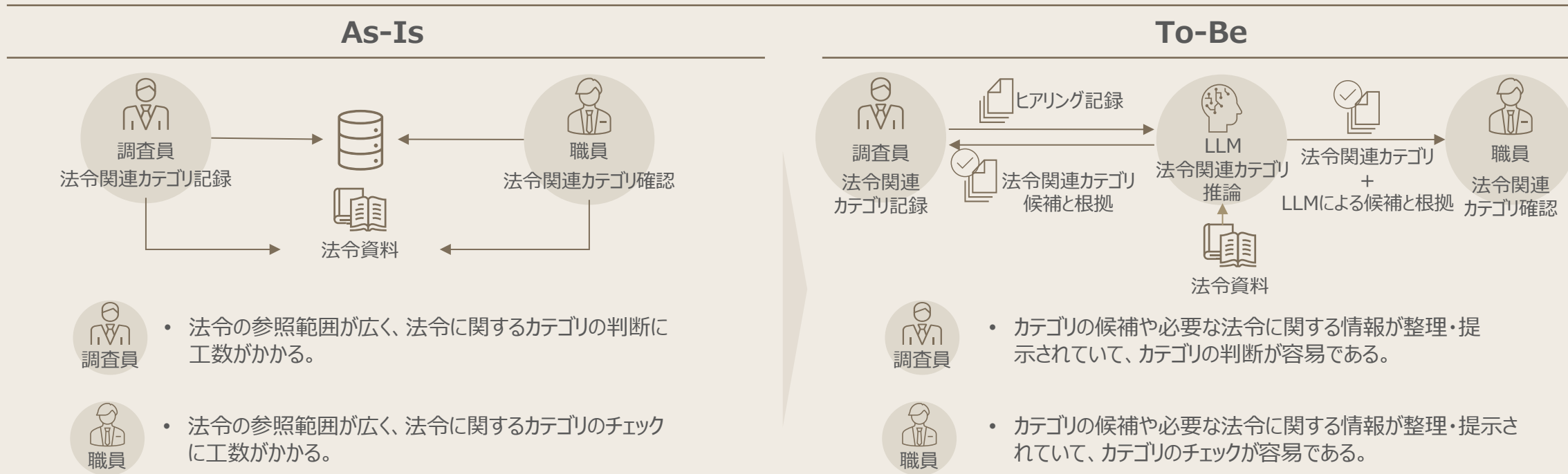
取引先基本情報

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 ④ 法令資料を参照するカテゴリ入力とチェックの効率化

- 法令関連のカテゴリの候補やその根拠、法令に関する情報が整理・提示されることで、入力・チェックの工数が削減される。

#### As-Is/To-Be案



#### 対象ヒアリング項目



#### テーマ別ヒアリング情報アプリ

価格交渉フォローアップ調査

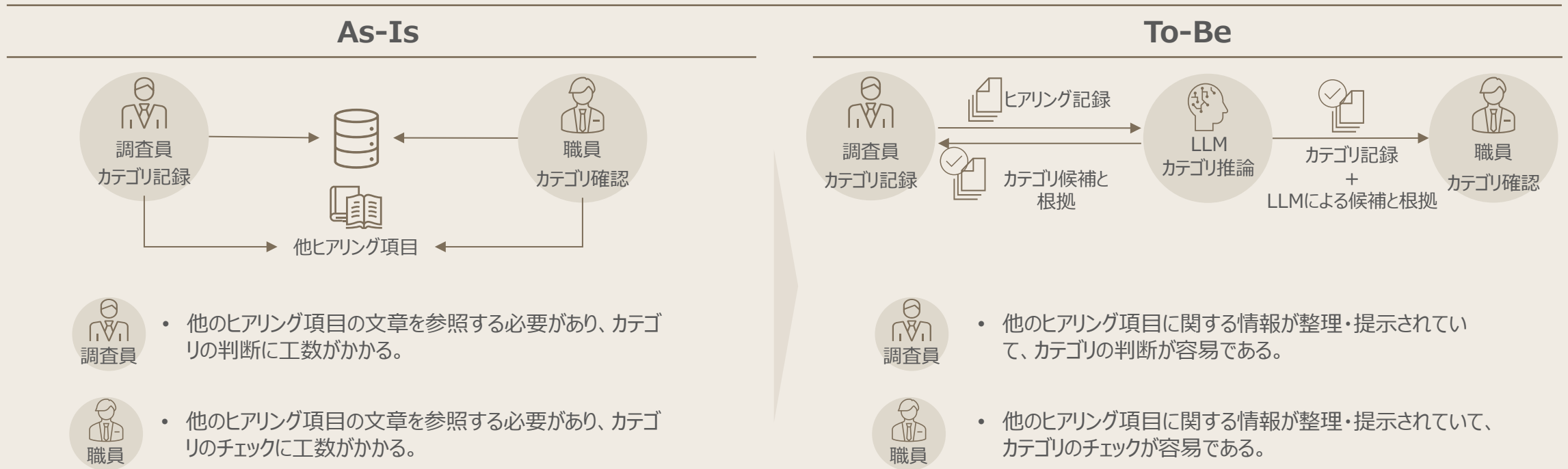
ヒアリング結果に対する所見等

### 3. 入力補助機能の有効性検証

#### 3.3. As-Is/To-Beの概要 ⑤法令資料を参照しないカテゴリ入力とチェックの効率化

- カテゴリの候補やその根拠が提示されることで、入力・チェックの工数が削減される。

#### As-Is/To-Be案



#### 対象ヒアリング項目



#### テーマ別ヒアリング情報アプリ

ヒアリング結果に対する所見等

### 3. 入力補助機能の有効性検証

#### 3.4. ユースケースの概要 ユースケースに対応するAs-Is/To-Be

- 5つのAs-Is/To-Beは、「A. フリー入力文章標準化」「B. 法令資料を参照しないカテゴリ値の入力・チェック効率化」「法令資料を参照するカテゴリ値の入力・チェック効率化」「フロー入力文章の整合性確認効率化」の4つのユースケースに集約される。

As-Is/To-Be名	対応ユースケース名	ユースケース概要
① 文章入力とチェックの効率化	A. フリー入力文章標準化	<ul style="list-style-type: none"><li>・ 調査員が入力した文章に対して、LLMが「文体や用語の統一」「文章の整理・整形」「内容別に文章を分類」「ヒアリング観点漏れ指摘」を行う。</li><li>・ これにより、調査員の文章入力と職員のチェック作業に対する工数削減と品質向上を実現する。</li></ul>
② 企業関連情報の 整合性チェックの効率化	D. フリー入力文章の整合性確認効率化	<ul style="list-style-type: none"><li>・ 調査員が入力した文章に対して、LLMが過去のヒアリング結果と入力内容に不整合がないかのチェックを行い、確認結果と根拠を提示する。</li><li>・ これにより、職員のチェック作業に対する工数削減と品質向上を実現する。</li></ul>
③ 業種カテゴリの 入力とチェックの効率化	B. 法令資料を参照しないカテゴリ値の入力・チェック効率化	<ul style="list-style-type: none"><li>・ LLMが調査員が入力した内容からカテゴリを推論し、推論結果とその根拠情報を提示する。</li><li>・ これにより、調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上を実現する。</li></ul>
④ 法令資料を参照する カテゴリ入力とチェックの効率化	C. 法令資料を参照するカテゴリ値の入力・チェック効率化	<ul style="list-style-type: none"><li>・ LLMが調査員が入力した内容からカテゴリを推論し、推論結果とその根拠情報や関連する法令情報を提示する。</li><li>・ これにより、調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上を実現する。</li></ul>
⑤ 法令資料を参照しない カテゴリ入力とチェックの効率化	B. 法令資料を参照しないカテゴリ値の入力・チェック効率化	<ul style="list-style-type: none"><li>・ LLMが調査員が入力した内容からカテゴリを推論し、推論結果とその根拠情報を提示する。</li><li>・ これにより、調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上を実現する。</li></ul>

### 3. 入力補助機能の有効性検証

#### 3.4. ユースケースの概要 A. フリー入力文章の標準化

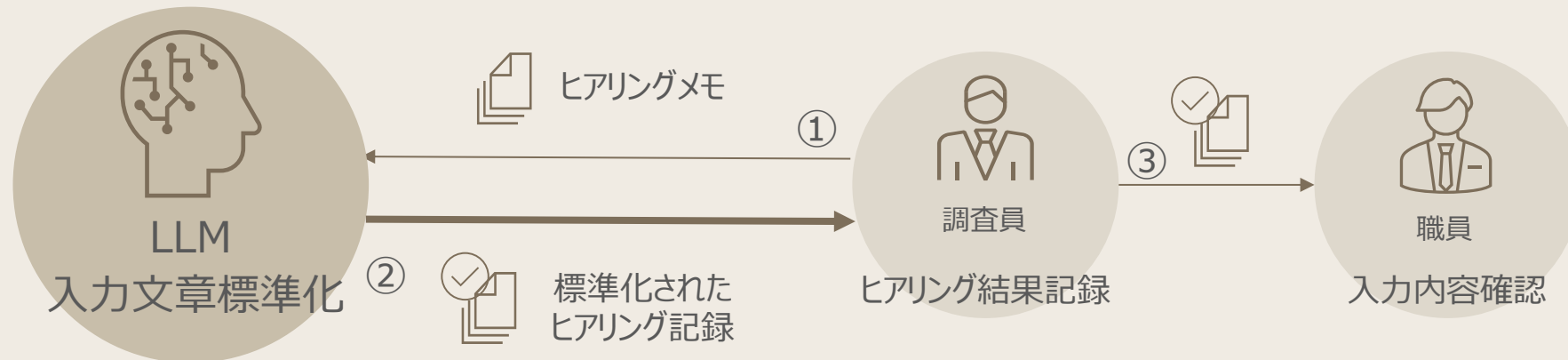
- 調査員が入力したヒアリング記録の文章をLLMが標準化し、標準化された文章を調査員に提示する。
- これにより調査員の文章入力と職員のチェック作業に対する工数削減と品質向上を実現する。

#### Before

- 調査員による文章整理、入力作業に工数がかかっている。
- 入力文章の品質が標準化されておらず、職員のチェック作業に工数がかかっている。

#### After

- 調査員が入力したヒアリング記録に対してLLMが標準化を行うことで、調査員の文章作成と職員のチェック作業に対する工数削減と品質向上が実現する。



### 3. 入力補助機能の有効性検証

#### 3.4. ユースケースの概要 B. 法令資料を参照しないカテゴリ値の入力・チェック効率化

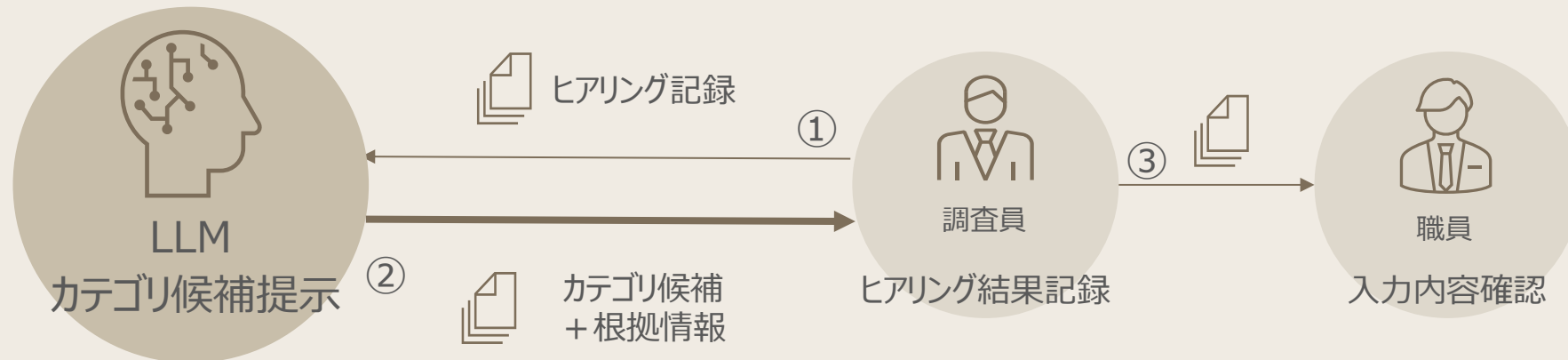
- 調査員が入力したヒアリング記録をもとにLLMがカテゴリ値を推論し、調査員に候補とその根拠を提示する。
- これにより調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上を実現する。

##### Before

- 調査員がルールや他のヒアリング項目を参照しながらカテゴリ値を入力する作業に工数がかかっている。
- 調査員によるカテゴリ値の入力の正確さが安定せず、職員のチェック作業に工数がかかっている。

##### After

- 調査員が入力したヒアリング記録をもとにLLMがカテゴリ値の候補とその根拠を提示することで、調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上が実現する。



### 3. 入力補助機能の有効性検証

#### 3.4. ユースケースの概要 C. 法令資料を参照するカテゴリ値の入力・チェック効率化

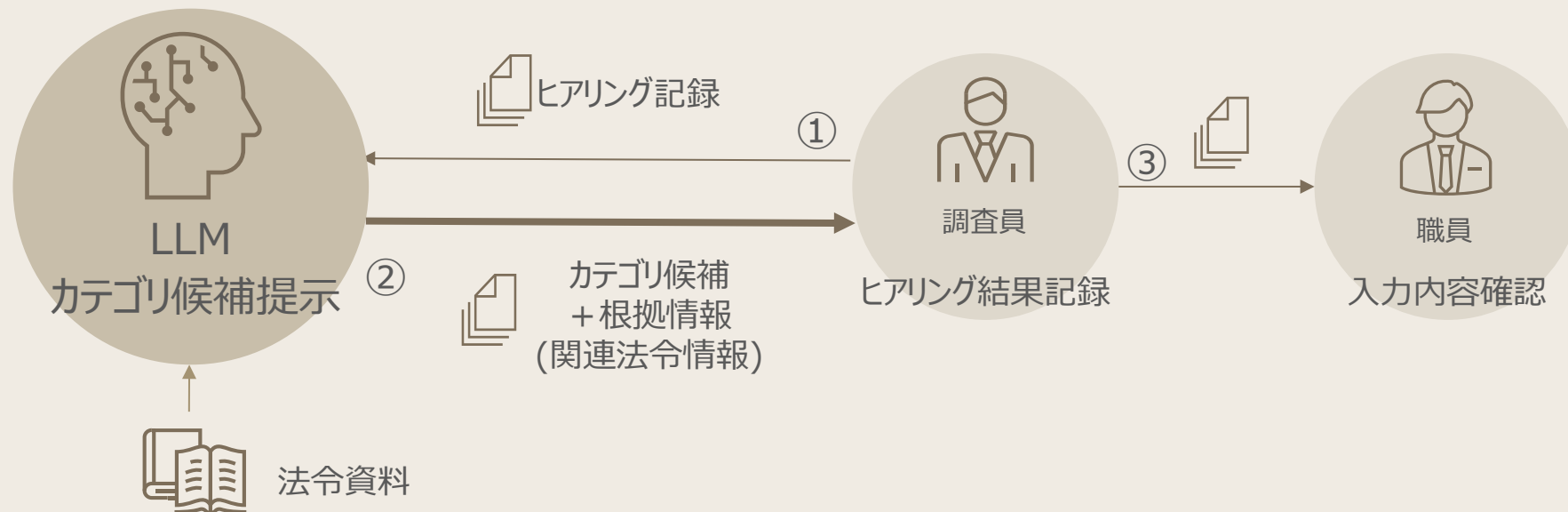
- 調査員が入力したヒアリング記録と法令資料をもとにLLMがカテゴリ値を推論し、調査員に候補とその根拠を提示する。
- これにより調査員のカテゴリ値入力と職員のチェックの工数削減と品質向上を実現する。

#### Before

- 調査員が法令を参照しながらカテゴリ値を入力する作業に工数がかかっている。
- 調査員によるカテゴリ値の入力の正確さが安定せず、職員のチェック作業にも工数がかかっている。

#### After

- 調査員が入力したヒアリング記録と法令資料をもとにLLMがカテゴリ値の候補とその根拠を提示することで、調査員のカテゴリ値入力と職員のチェック作業に対する工数削減と品質向上が実現する。



### 3. 入力補助機能の有効性検証

#### 3.4. ユースケースの概要 D. フリー入力文章の整合性確認効率化

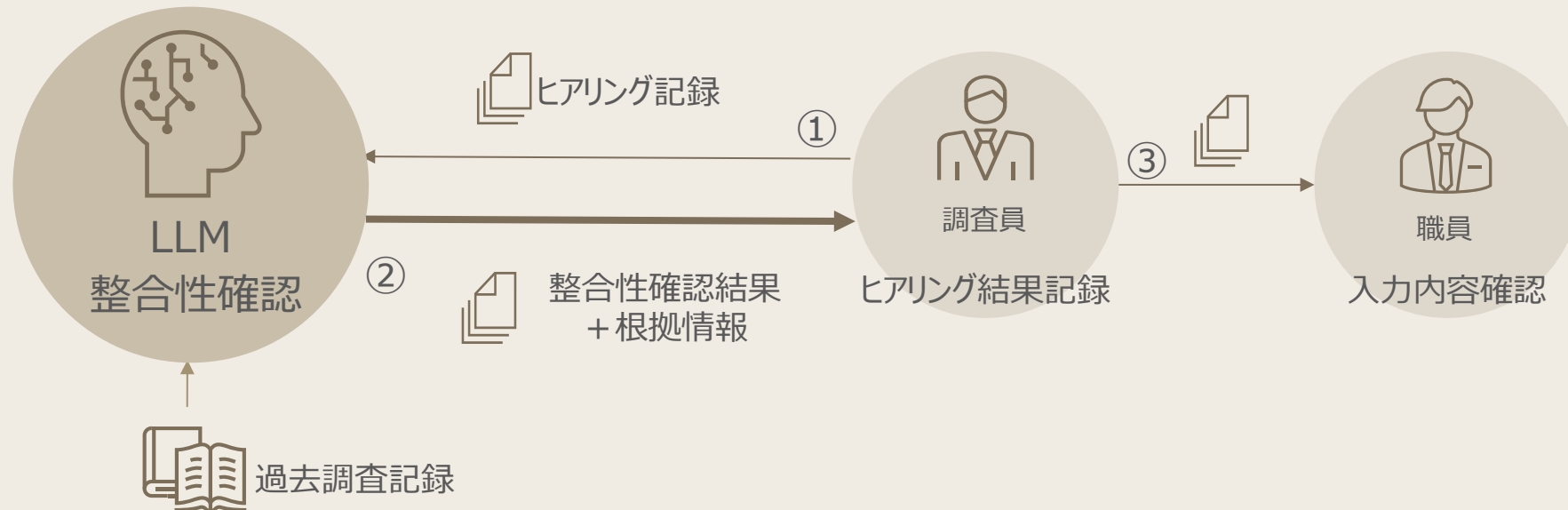
- 調査員が入力したヒアリング記録の文章に対して、過去のヒアリング結果から内容に誤りの可能性がないかの確認を行い、その確認結果と根拠を提示する。
- これによりヒアリング記録の品質向上、職員のチェック作業に対する工数削減を実現する。

#### Before

- 入力文章の内容に誤りがないかのチェック作業に工数がかかっている。

#### After

- 調査員が入力したヒアリング記録に対してLLMが内容の整合性チェックを行うことで、職員のチェック作業に対する工数削減が実現する。



### 3. 入力補助機能の有効性検証

#### 3.5. 各ユースケースの対象ヒアリング項目

- 各ユースケースでの効果が見込まれる項目を対象にした。

ユースケース名	アプリ	対象ヒアリング項目分類
A. フリー入力文章標準化	ヒアリング先基本情報	ヒアリング先中小企業基本情報 ----- 自社が委託事業者となる場合の取引の状況 ----- 備考
	取引先情報	取引先基本情報 -----
	テーマ別ヒアリング情報	価格交渉フォローアップ調査 ----- ヒアリング結果に対する所見等 -----
B. 法令資料を参照しない カテゴリ値の入力・チェック効率化	ヒアリング先基本情報	ヒアリング先中小企業基本情報 -----
	取引先情報	取引先基本情報 -----
	テーマ別ヒアリング情報	ヒアリング結果に対する所見等 -----
C. 法令資料を参照する カテゴリ値の入力・チェック効率化	テーマ別ヒアリング情報	価格交渉フォローアップ調査 ----- ヒアリング結果に対する所見等 -----
D. フロー入力文章の 整合性確認効率化	ヒアリング先基本情報	ヒアリング先中小企業基本情報 -----
	取引先情報	取引先基本情報 -----

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 A. フリー入力文章の標準化（1/3）

- フリー入力文章の標準化の7つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
文体統一	敬体・文体/ 体言・用言止めの統一	文体統一による 可読性	LLMが行った文体統一により文章が読みやすくなったか。	文体変化による「文章全体の読みやすさの変化（素点）」と「読みにくい箇所数（減点）」からスコアリング
		文体統一による 信頼性	LLMが適切に文体統一を行っているか。	LLMが行った文体統一が適切だった割合を適合率※1で評価
		文体統一による 網羅性	LLMが網羅的に文体統一を行っているか。	LLMが文体統一をどれだけ漏れなく行えたかを再現率※2で評価
用語統一	同義語や略語の統一	用語統一による 可読性	LLMが行った用語統一により文章が読みやすくなったか。	文体変化による「文章全体の読みやすさの変化（素点）」と「読みにくい箇所数（減点）」からスコアリング
		用語統一による 信頼性	LLMが適切に用語統一を行っているか。	LLMが行った文体統一が適切だった割合を適合率※1で評価
		用語統一による 網羅性	LLMが網羅的に用語統一を行っているか。	LLMが文体統一をどれだけ漏れなく行えたかを再現率※2で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)  
 ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 A. フリー入力文章の標準化 (2/3)

- フリー入力文章の標準化の7つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
曖昧表現の 明確化	「適宜」「思う」などの 曖昧表現を明確化	一次情報 特性保持	LLMが一次情報の特性を失わないように 曖昧表現を明確化できているか	曖昧表現の明確化による「文章全体の一次情報特性の 保持具合（素点）」と「一次情報特性を失っている箇所 数（減点）」からスコアリング
		意味保持性	LLMが意味を保持したまま曖昧表現を明 確化できているか	曖昧表現の明確化による「文章全体の意味の保持具合 （素点）」と「意味が変化している箇所数（減点）」から スコアリング
文章種別分類	文章を内容別に分類	種別分類の 信頼性	LLMが適切に文章を振り分けられている か。	LLMが行った文章種別分類が適切だった割合を適合率 ※1で評価
		種別分類の 網羅性	LLMが網羅的に文章を振り分けられている か。	LLMが文章種別分類統一をどれだけ漏れなく行えたかを再 現率※2で評価
ヒアリング観点 指摘	ヒアリング観点の 漏れを確認	ヒアリング観点 指摘信頼性	LLMが適切にヒアリング観点漏れを指摘で きているか。	LLMがヒアリング観点指摘が適切だった割合を適合率※1 で評価
		ヒアリング観点 指摘網羅性	LLMが網羅的にヒアリング観点漏れを指摘 できているか。	LLMがヒアリング観点指摘をどれだけ漏れなく行えたかを再 現率※2で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)  
 ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 A. フリー入力文章の標準化 (3/3)

- フリー入力文章の標準化の7つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
表現の簡潔化	冗長表現や重複を整理・削除	簡潔化による可読性	LLMが文章を簡潔にすることにより読みやすいになっているか	曖昧表現の明確化による「文章全体の一次情報特性の保持具合（素点）」と「一次情報特性を失っている箇所数（減点）」からスコアリング
		簡潔化による意味保持性	LLMが意味を保持したまま表現を簡潔化できているか	曖昧表現の明確化による「文章全体の意味の保持具合（素点）」と「意味が変化している箇所数（減点）」からスコアリング
標準化の過程の記録	書き換えの概要・根拠を明記	記録の理解容易性	LLMの標準化の過程の記録が理解しやすいか。	「標準化の過程の記録の理解しやすさ（素点）」と「理解しにくい箇所数（減点）」からスコアリング
		記録の信頼性	LLMが適切に標準化の過程を記録できているか。	LLMが行った標準化の過程の記録のうち適切だった割合を適合率※1で評価
		記録の網羅性	LLMが網羅的に標準化の過程を記録できているか。	LLMが標準化の過程の記録をどれだけ漏れなく行えたかを再現率※2で評価
		フォーマットの正確性	LLMが標準化の過程を記録を正しいフォーマットできているか。	LLMが行った標準化の過程のうち正しいフォーマットでの出力である割合を正解率※3で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)
- ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)
- ※3. 正確率：LLMが正しいフォーマットで出力できた割合(LLMの出力の正確性)

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 B. 法令資料を参照しないカテゴリ値の入力・チェック効率化

- 法令資料を参照しないカテゴリ値の入力・チェック効率化の2つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
カテゴリ値候補提示	カテゴリ値の候補を提示	カテゴリ値候補提示の信頼性	LLMが適切にカテゴリ値の候補提示をできているか。	LLMが行ったカテゴリ値候補提示のうち適切だった割合を適合率※1で評価
		カテゴリ値候補提示の網羅性	LLMが網羅的にカテゴリ値の候補提示をできているか。	LLMがカテゴリ値候補提示をどれだけ漏れなく行えたかを再現率※2で評価
		カテゴリ値制約遵守の正確性	LLMが定義されているカテゴリリスト内からカテゴリを提示できているか。	LLMが提示したカテゴリ値候補が定義されたカテゴリリストに含まれる割合を正解率※3で評価
根拠の記録	そのカテゴリを提示した根拠を記録	記録の理解容易性	LLMの根拠の記録が理解しやすいか。	「根拠の記録の理解しやすさ（素点）」と「理解しにくい箇所数（減点）」からスコアリング
		根拠の正確性	LLMが正確な根拠を記録できているか。	「根拠の記録全体の正確さ（素点）」と「誤っている箇所数（減点）」からスコアリング
		フォーマットの正確性	LLMが根拠の記録を正しいフォーマットできているか。	LLMが行った根拠の記録のうち正しいフォーマットでの出力である割合を正解率※3で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)
- ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)
- ※3. 正確率：LLMが正しいフォーマットで出力できた割合(LLMの出力の正確性)

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 C. 法令資料を参照するカテゴリ値の入力・チェック効率化

- 法令資料を参照するカテゴリ値の入力・チェック効率化の2つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
カテゴリ値候補提示	法令資料を参照してカテゴリ値の候補を提示	カテゴリ値候補提示の信頼性	LLMが適切にカテゴリ値の候補提示をできているか。	LLMが行ったカテゴリ値候補提示のうち適切だった割合を適合率※1で評価
		カテゴリ値候補提示の網羅性	LLMが網羅的にカテゴリ値の候補提示をできているか。	LLMがカテゴリ値候補提示をどれだけ漏れなく行えたかを再現率※2で評価
		カテゴリ値制約遵守の正確性	LLMが定義されているカテゴリリスト内からカテゴリを提示できているか。	LLMが提示したカテゴリ値候補が定義されたカテゴリリストに含まれる割合を正解率※3で評価
根拠の記録	そのカテゴリを提示した根拠を記録	記録の理解容易性	LLMの根拠の記録が理解しやすいか。	「根拠の記録の理解しやすさ（素点）」と「理解しにくい箇所数（減点）」からスコアリング
		根拠の正確性	LLMが正確な根拠を記録できているか。	「根拠の記録全体の正確さ（素点）」と「誤っている箇所数（減点）」からスコアリング
		フォーマットの正確性	LLMが根拠の記録を正しいフォーマットできているか。	LLMが行った根拠の記録のうち正しいフォーマットでの出力である割合を正解率※3で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)
- ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)
- ※3. 正確率：LLMが正しいフォーマットで出力できた割合(LLMの出力の正確性)

### 3. 入力補助機能の有効性検証

#### 3.6. 検証における評価観点と評価方法 D. フリー入力文章の整合性確認効率化

- フリー入力文章のチェック効率化の2つの機能を検証・評価した。
- 適合率・再現率・正解率を用いた定量評価と、スコアリングを用いた定性評価により評価をした。

機能		評価		
機能名	機能概要	評価項目名	評価項目概要	評価方法
整合性確認 結果提示	過去のヒアリング結果や 調査結果項目を 参照し、文章内容の 整合性を確認	整合性確認 結果の信頼性	LLMが適切に整合性確認をできているか。	LLMが行った整合性確認のうち適切だった割合を適合率 ※1で評価
		整合性確認 結果の網羅性	LLMが網羅的に整合性確認をできているか。	LLMが整合性確認をどれだけ漏れなく行えたかを再現率 ※2で評価
根拠の記録	その整合性確認結果を 提示した根拠を記録	記録の 理解容易性	LLMの根拠の記録が理解しやすいか。	「根拠の記録の理解しやすさ（素点）」と「理解しにくい箇所数（減点）」からスコアリング
		根拠の正確性	LLMが正確な根拠を記録できているか。	「根拠の記録全体の正確さ（素点）」と「誤っている箇所数（減点）」からスコアリング
		フォーマットの 正確性	LLMが根拠の記録を正しいフォーマットできているか。	LLMが行った根拠の記録のうち正しいフォーマットでの出力である割合を正解率※3で評価

- ※1. 適合率：LLMが標準化した箇所うち、実際に正しく標準化できていた割合(LLMの判定の信頼性)
- ※2. 再現率：標準化すべき箇所のうち、LLMが正しく標準化することができた割合(LLMの判定の網羅性)
- ※3. 正確率：LLMが正しいフォーマットで出力できた割合(LLMの出力の正確性)

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 評価結果サマリ

- すべてのユースケースの機能において、定量評価80%以上、定性評価8.5点以上であった。
- これは運用に耐える精度とスコアと考えられる。

ユースケース名	機能	定量評価結果			定性評価結果
		適合率	再現率	正解率	スコアリング
A. フリー入力文章標準化	文体統一	100%	87%	-	10/10
	用語統一	91%	100%	-	10/10
	曖昧表現の明確化	調査員の記録のニュアンスが失われる可能性があり、また、可読性向上はわずかなため機能削除			
	文章種別分類	85%	88%	-	-
	ヒアリング観点指摘	95%	100%	-	-
	表現の簡潔化	-	-	-	8.8/10
	標準化の過程の記録	84%	92%	100%	10/10
B. 法令資料を参照しない カテゴリ値の入力・チェック効率化	カテゴリ値候補提示	86%	86%	100%	9.0/10
	根拠の記録	-	-	100%	9.6/10
C. 法令資料を参照する カテゴリ値の入力・チェック効率化	カテゴリ値候補提示	91%	93%	100%	9.5/10
	根拠の記録	-	-	100%	10/10
D. フリー入力文章の整合性確認効率化	整合性確認結果提示	80%	80%	-	9.4/10
	根拠の記録	-	-	100%	9.1/10

適合率：LLMが行った標準化・カテゴリ分類・整合性判定のうち、どれだけ正確に行えたかをはかる指標  
 再現率：LLMが対象となる標準化・カテゴリ分類・整合性判定を、どれだけ漏れなく行えたかをはかる指標  
 正解率：LLMがどれだけ正しい出力できたかをはかる指標  
 ※適合率・再現率はマクロ平均（各機能の精度の平均値）を記載

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 A. フリー入力文章の標準化（1/2）

- ユースケースAの機能の評価結果は、定量評価80%以上、定性評価8.5点以上と運用に耐えうる精度とスコアであると考えられる。
- ただし、「曖昧表現の明確化」は調査員の記録のニュアンスが失われる可能性があり、また、可読性向上もわずかであるため実装しない。

機能名	評価項目名	評価項目概要	評価結果		
			定量評価	定性評価	評価指標
文体統一	文体統一による可読性	LLMが行った文体統一により文章が読み易くなったか。	-	10/10	スコアリング
	文体統一による信頼性	LLMが適切に文体統一を行っているか。	100%	-	適合率
	文体統一による網羅性	LLMが網羅的に文体統一を行っているか。	87%	-	再現率
用語統一	用語統一による可読性	LLMが行った用語統一により文章が読み易くなったか。	-	10/10	スコアリング
	用語統一による信頼性	LLMが適切に用語統一を行っているか。	91%	-	適合率
	用語統一による網羅性	LLMが網羅的に用語統一を行っているか。	100%	-	再現率
曖昧表現の明確化	一次情報特性保持	LLMが一次情報の特性を失わないように曖昧表現を明確化できているか。	調査員の記録のニュアンスが失われる可能性があり、また、可読性向上もわずかであるため実装しない		
	意味保持性	LLMが意味を保持したまま曖昧表現を明確化できているか。			

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 A. フリー入力文章の標準化 (2/2)

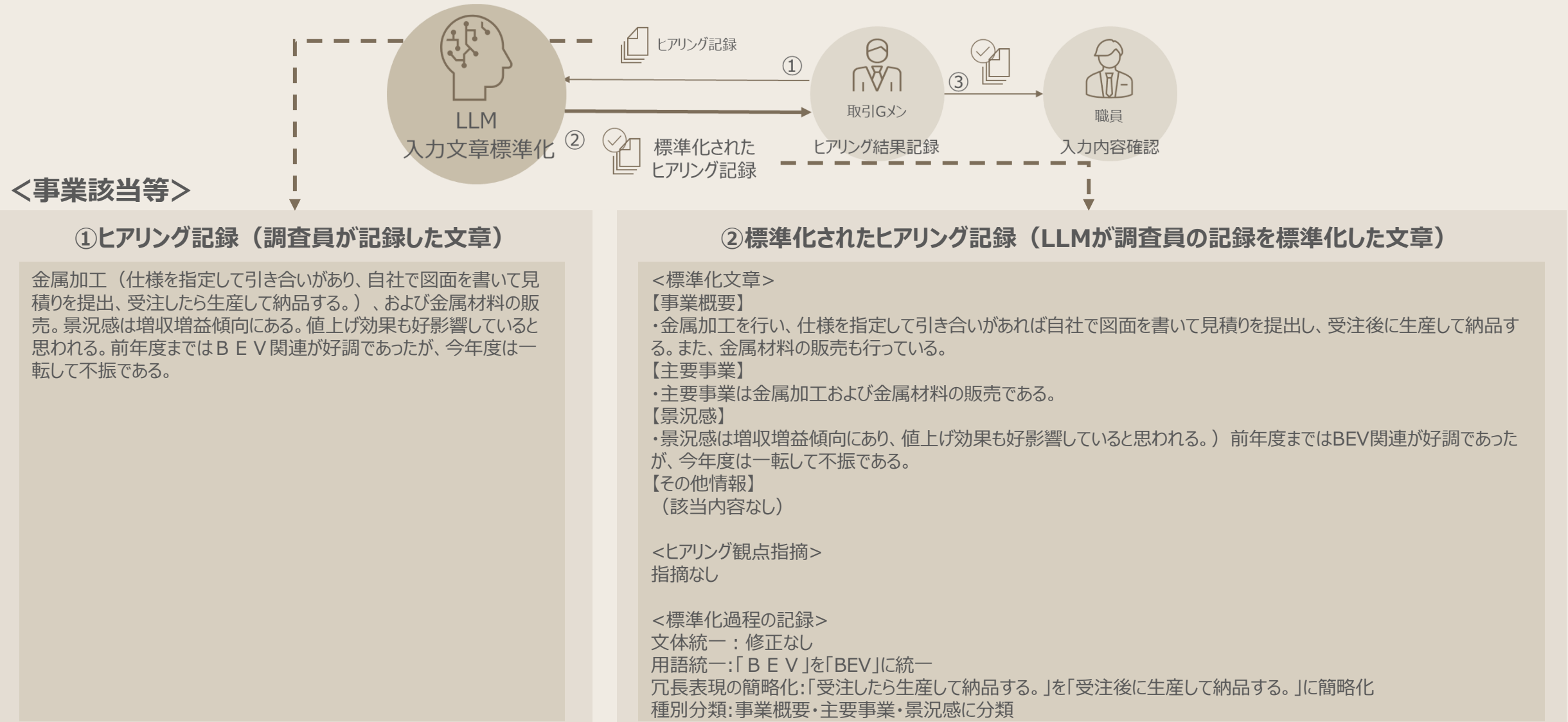
- ユースケースAの機能の評価結果は、定量評価80%以上、定性評価8.5点以上と運用に耐えうる精度とスコアであると考えられる。
- ただし、「曖昧表現の明確化」は調査員の記録のニュアンスが失われる可能性があり、また、可読性向上もわずかであるため実装しない。

機能名	評価項目名	評価項目概要	評価結果		
			定量評価	定性評価	評価指標
文章種別分類	種別分類の信頼性	LLMが適切に文章を振り分けているか。	85%	-	適合率
	種別分類の網羅性	LLMが網羅的に文章を振り分けているか。	88%	-	再現率
ヒアリング観点指摘	ヒアリング観点指摘信頼性	LLMが適切にヒアリング観点漏れを指摘できているか。	95%	-	適合率
	ヒアリング観点指摘網羅性	LLMが網羅的にヒアリング観点漏れを指摘できているか。	100%	-	再現率
表現の簡潔化	簡潔化による可読性	LLMが文章を簡潔にすることにより読み易くなっているか	-	8.8/10	スコアリング
	簡潔化による意味保持性	LLMが意味を保持したまま表現を簡潔化できているか	-	8.8/10	スコアリング
標準化の過程の記録	記録の理解容易性	LLMによる標準化の過程の記録が理解しやすいか。	-	10/10	スコアリング
	記録の信頼性	LLMが適切に標準化の過程を記録できているか。	84%	-	適合率
	記録の網羅性	LLMが網羅的に標準化の過程を記録できているか。	92%	-	再現率
	フォーマットの正確性	LLMが標準化の過程を記録を正しいフォーマットできているか。	100%	-	正解率

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 A. フリー入力文章の標準化 具体例

- ①のような調査員の入力文章に対して、LLMが②のような標準化された文章とヒアリング観点のチェック結果、標準化の過程を提示する。



### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 B. 法令資料を参照しないカテゴリ値の入力・チェック効率化

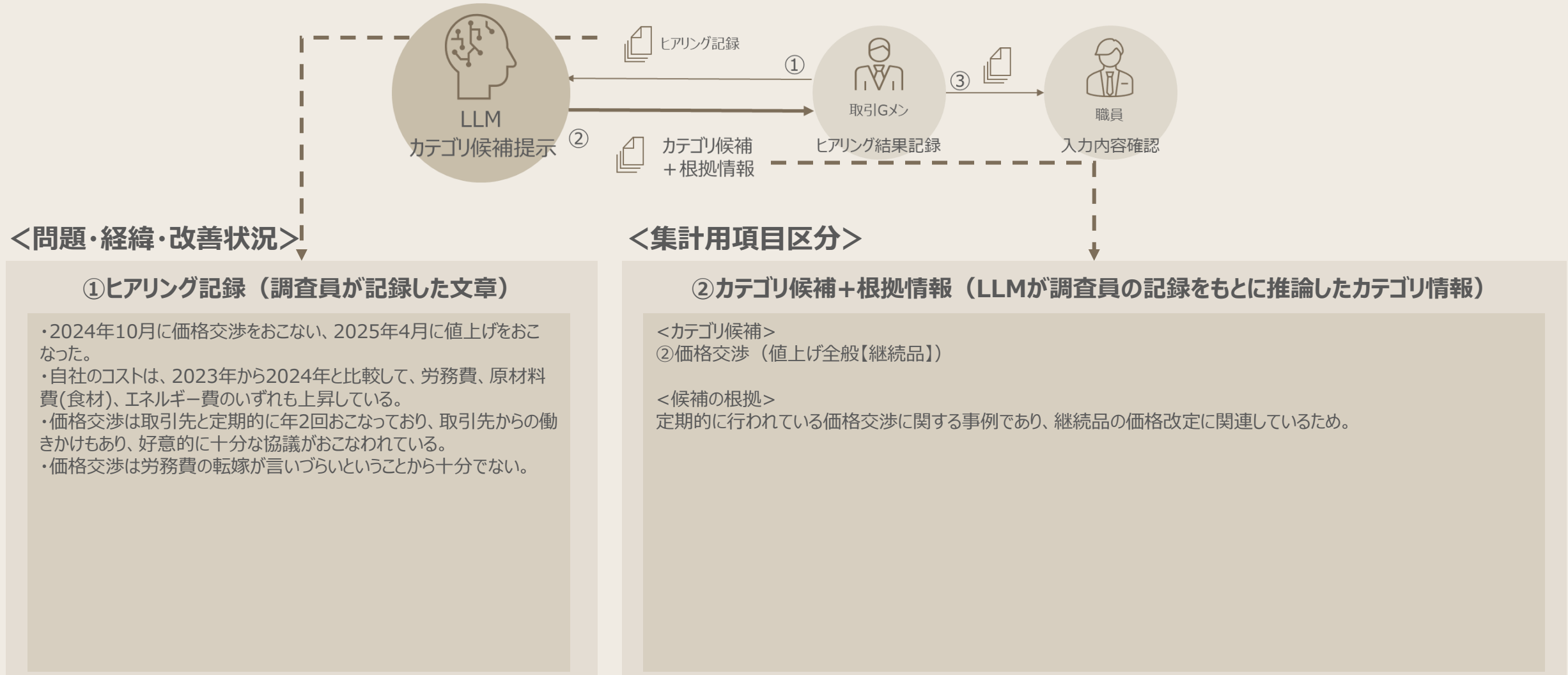
- ユースケースBの機能の評価結果は、定量評価75%以上、定性評価9.0点以上と運用に耐えうる精度とスコアであると考えられる。

機能名	評価項目名	評価項目概要	評価結果		
			定量評価	定性評価	評価指標
カテゴリ値候補提示	カテゴリ値候補提示の信頼性	LLMが適切にカテゴリ値の候補提示をしているか。	87%	-	適合率
	カテゴリ値候補提示の網羅性	LLMが網羅的にカテゴリ値の候補提示をしているか。	86%	-	再現率
	カテゴリ値制約遵守の正確性	LLMは定義されているカテゴリリスト内からカテゴリを提示できているか。	100%	-	正解率
根拠の記録	記録の理解容易性	LLMの根拠の記録が理解しやすいか。	-	9.0/10	スコアリング
	根拠の正確性	LLMが正確な根拠を記録できているか。	-	9.6/10	スコアリング
	フォーマットの正確性	LLMが根拠の記録を正しいフォーマットできているか。	100%	-	正解率

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 B. 法令資料を参照しないカテゴリ値の入力・チェック効率化 具体例

- ①のような調査員が記録した文章をもとに、LLMが②カテゴリ候補 + 根拠情報を提示する。



### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 C. 法令資料を参照するカテゴリ値の入力・チェック効率化

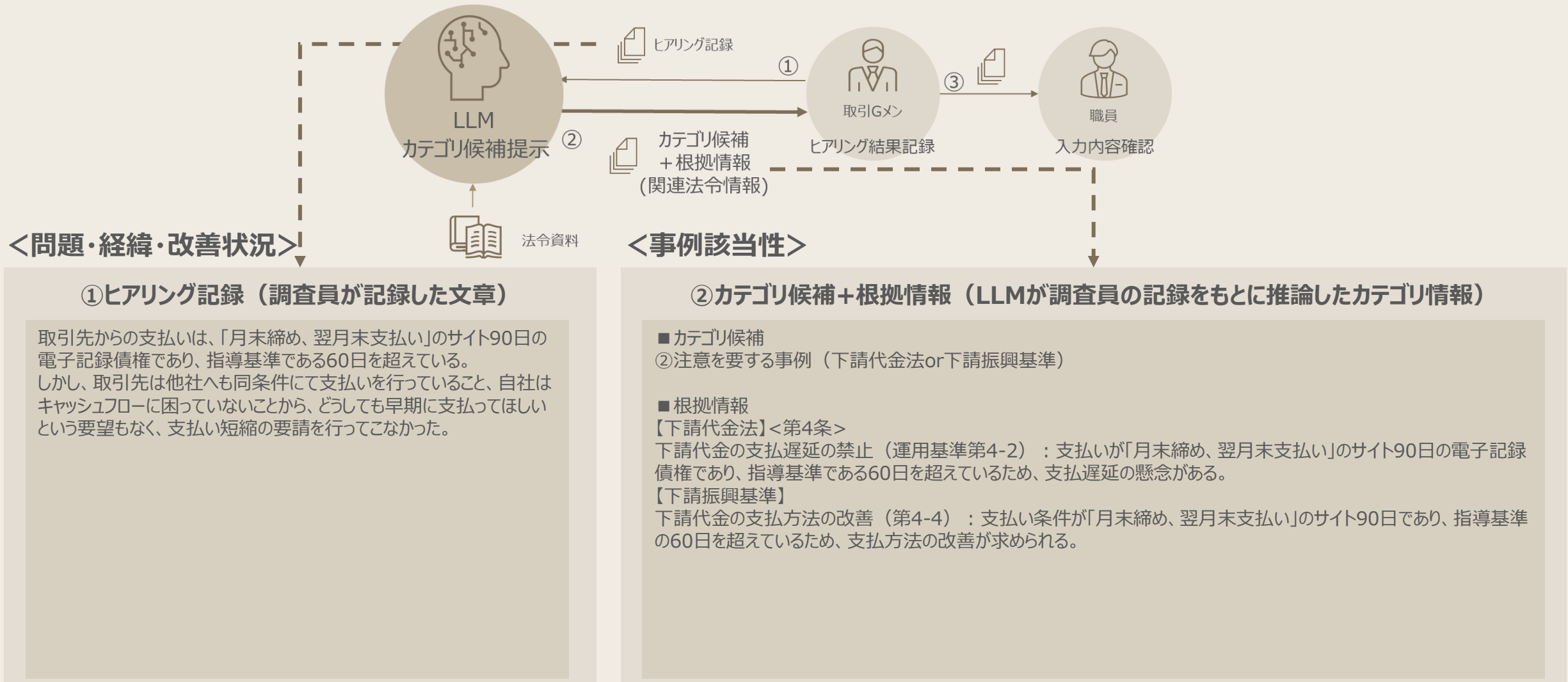
- ユースケースCの機能の評価結果は、定量評価85%以上、定性評価9.0点以上と運用に耐えうる精度とスコアであると考えられる。

機能名	評価項目名	評価項目概要	評価結果		
			定量評価	定性評価	評価指標
カテゴリ値候補提示	カテゴリ値候補提示の信頼性	LLMが適切にカテゴリ値の候補提示をしているか。	91%	-	適合率
	カテゴリ値候補提示の網羅性	LLMが網羅的にカテゴリ値の候補提示をしているか。	93%	-	再現率
	カテゴリ値制約遵守の正確性	LLMは定義されているカテゴリリスト内からカテゴリを提示できているか。	100%	-	正解率
根拠の記録	記録の理解容易性	LLMの根拠の記録が理解しやすいか。	-	9.5/10	スコアリング
	根拠の正確性	LLMが正確な根拠を記録できているか。	-	10/10	スコアリング
	フォーマットの正確性	LLMが根拠の記録を正しいフォーマットできているか。	100%	-	正解率

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 C. 法令資料を参照するカテゴリ値の入力・チェック効率化 具体例

- ①のような調査員が記録した文章をもとに、LLMが②カテゴリ候補 + 根拠情報を提示する。



### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 D. フリー入力文章の整合性確認効率化

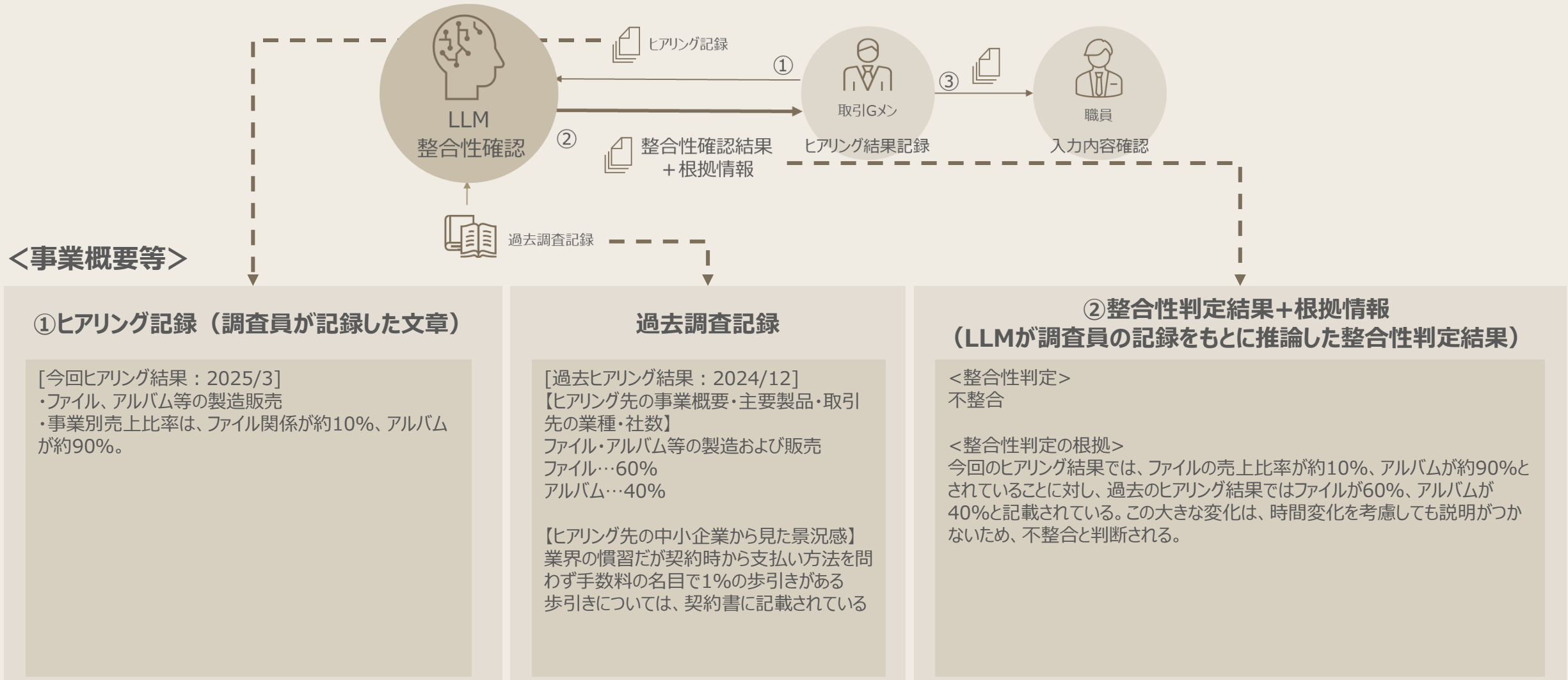
- ユースケースDの機能の評価結果は、定量評価80%以上、定性評価9.0点以上と運用に耐えうる精度とスコアであると考えられる。

機能名	評価項目名	評価項目概要	評価結果		
			定量評価	定性評価	評価指標
整合性確認 結果提示	整合性確認 結果の信頼性	LLMが適切に整合性確認をできているか。	80%	-	適合率
	整合性確認 結果の網羅性	LLMが適切に整合性確認をできているか。	80%	-	再現率
根拠の記録	記録の 理解容易性	LLMによる根拠の記録が理解しやすいか。	-	9.4/10	スコアリング
	根拠の正確性	LLMが正確な根拠を記録できているか。	-	9.1/10	スコアリング
	フォーマットの 正確性	LLMによる根拠の記録は正しいフォーマットで出力できているか。	100%	-	正解率

### 3. 入力補助機能の有効性検証

#### 3.7. 評価結果 D. フリー入力文章の整合性確認効率化 具体例

- ①のような調査員が記録した文章をもとに、LLMが②のような整合性確認結果と根拠情報を提示する。



## 4. 入力補助機能の環境構築および検証

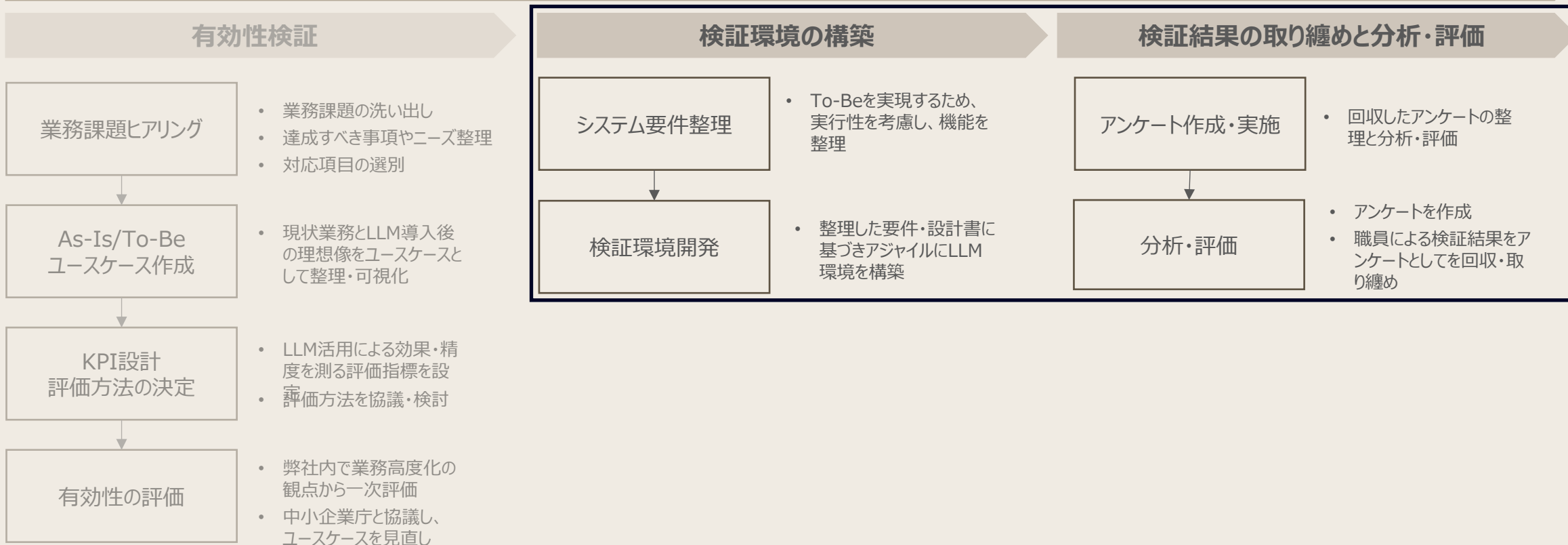
- 4.1. 検証の実施方針
- 4.2. 検証環境の要件
- 4.3. 検証環境の構成・仕様
- 4.4. 基本機能と利用イメージ
- 4.5. 検証方法
- 4.6. 検証結果
- 4.7. 今後対応すべき課題

## 4. 入力補助機能の環境構築および検証

### 4.1. 検証の実施方針

- 業務ヒアリングを通じて定義した将来像（To-Be）およびユースケースを基に検証環境を構築し、職員による検証結果を分析することで、取引調査員の業務高度化への有効性を評価した。

#### 入力補助機能の検証プロセス



## 4. 入力補助機能の環境構築および検証

### 4.2. 検証環境の要件

- 本検証環境で具備した機能要件・非機能要件を一覧に整理した。

機能要件	詳細	実現方法
ログイン機能	予め設定されたアカウントでのみ、本サービスを利用可能となるよう、主体認証を備えたログイン機能を設けること。なお、中小企業庁職員がパスワードを失念した場合に初期化または再発行が可能である必要があるため、当該機能も合わせて有すること。	kintoneからLLMモデル呼び出しにおいて、kintone側でのユーザー制御を実施。APIはIP制限とキーによる保護を実施
コンテンツ生成機能	中小企業庁職員が登録したデータによりLLMを適切に利用することで、以下の項目の実現性を検討する。 ① ヒアリング結果入力補助 中小企業庁職員が調査結果を入力後に、LLMが入力内容の不明確な点や抜け漏れがないかをサポートすることで、入力内容の質向上を行う。 ② 事例判定補助 調査結果および業種の情報から、対象となる法律・ガイドライン・行動計画等を参照し、事例判定情報の付与を行う。 ③ 業種別分析業種別の調査結果より事例を集約し、業種別の傾向分析を実現するための手法を確立する。 ④ 法令/業種の誤り判定 ヒアリング内容から参照すべき法令や業種レベルが妥当であるかの判断がLLMで実施可能かを検討する。	入力補助機能（既存kintone同等環境とAzureの連携）および業種別分析機能（Azure環境のみ）の2機能を提供
利用状況分析機能	中小企業庁職員による本サービスの利用状況を確認、分析が可能となる機能を提供すること。具体的には、トレーサビリティを確保する目的を前提にアクセスログを取得し、別途分析可能な記録を取得する。なお、アクセスログのフォーマットはユーザー単位とし、日時、アクション、対象範囲等を確認できるようにすること。	Azureにおけるユーザー単位でのアクセスログ取得は検証段階では実施しないが、kintoneにおけるアクセスログは可能
非機能要件	詳細	実現方法
キャパシティ要件	中小企業庁職員の15名程度の登録および利用を想定すること。なお、本事業期間中にコンテンツ生成機能を用いた入出力トークン数の合計は5億トークン程度を想定すること。上記トークン数は、事業期間中に中小企業庁職員が可能な限り本環境を利用できるよう、消費トークン数を低減する工夫も入れ込んで実装すること。	ユーザー数10人以上の登録、および、利用できる形で運用
利用環境要件	本サービスの利用環境は、当省基盤情報システムにて提供されるセキュアPC、および同システムネットワークからの接続を想定し、サービスへの接続を最小限とすること。また、本システムへアクセスするブラウザについては以下を想定すること。 HP Sure Click Secure Browser	職員PCからアプリへのアクセス、アプリ利用可能
ユーザビリティ要件	政府統一指針への対応、画面構成、操作方法の分かりやすさ、指示や状態の分かりやすさ、エラーの防止と処理、ヘルプ	デザインシステムを踏襲したUI開発実施
アクセシビリティ要件	日本産業規格 JIS X 8341 シリーズ、「みんなの公共サイト運用ガイドライン」（総務省）、経済産業省ウェブアクセシビリティ方針を参考に、アクセシビリティを確保した設計・開発。指示や状態の分かりやすさ	中小企業庁担当職員と協議しながらUI開発実施
可用性・継続性要件	メンテナンス時間を除く24時間365日の稼働が前提であることとするが、中小企業庁職員による利用期間、タイムスケジュールについては事業開始後、担当職員と協議の上、稼働時間を設定すること	本事業安定稼働（リリース時除く）
性能要件	本事業では要件として提示しないが、中小企業庁職員が最低限利用可能となるようにすること。	中小企業庁職員利用可能な環境を提供
セキュリティ要件	本事業で扱う情報は、機密性2までの情報に限る。中小企業庁職員が誤って機密性3のデータを登録しないよう、本環境においてWeb画面を設ける場合においては注意喚起を促すメッセージを表示する等、対策を講じること。当サービスを利用することで生じる入力情報、出力情報のすべては、事業者の管理する環境内のみ蓄積し、外部への流出が発生しないよう十分な設定を行うとともに、適切に管理されていることをサービス提供期間中は確認すること。また、事業期間外においては当該情報を適切に破棄すること。加えて、利用者が中小企業庁職員に限定されている点を鑑み、適切なアクセス制御を行うことも検討すること。	機密性2までの情報のみ扱う 誤入力防止の注意喚起 入出力情報は環境内で管理、期間終了後は破棄。アクセス制御は協議の上、適切に実施
運用要件	受託者は中小企業庁職員の利用において消費したキャパシティを日次で監視し、翌営業日にキャパシティ使用状況を担当課室に報告すること。また、中小企業庁職員が有益に利用可能なよう中小企業庁職員のニーズに対応するとともに、環境を利用する際に受託者の提供範囲外で発生した課題に対して積極的に解決策を検討すること。	消費キャパシティの監視は未実施。検証を滞りなく進めるため発生した課題に対し迅速に対応
保守要件	サービス提供不可となった場合は迅速な復旧に努めること。また、利用する製品・サービスに脆弱性を検出した場合は、速やかに対処すること	随時トラブルシューティング実施

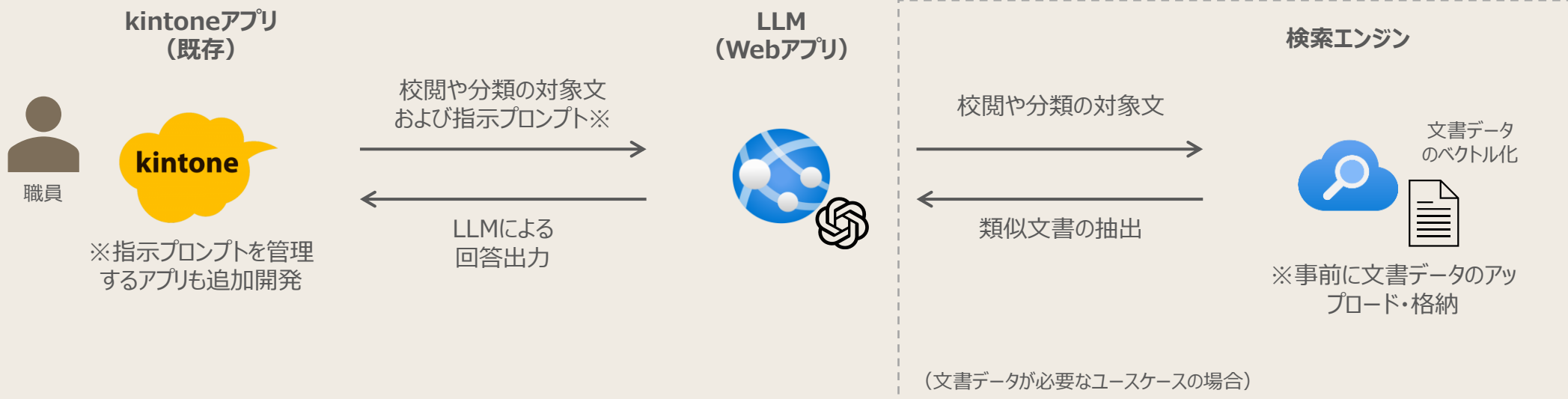
※第5章の業種別分析機能においても上記要件にて実施

## 4. 入力補助機能の環境構築および検証

### 4.3. 検証環境の構成・仕様

- 既存のkintone画面に、校閲・分類入力を支援するためのLLMや検索エンジンを組み込み、検証環境を準備した。

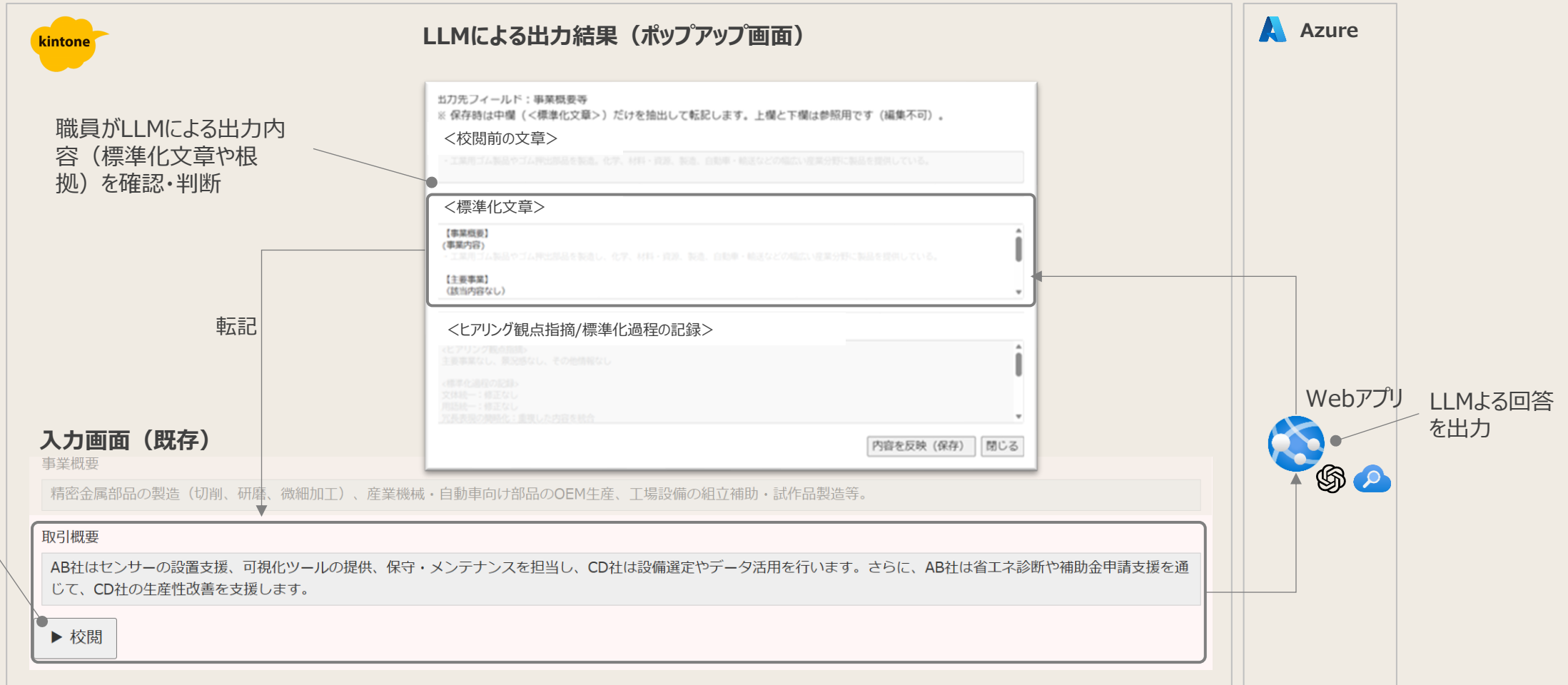
#### 入力補助機能



## 4. 入力補助機能の環境構築および検証

### 4.4. 基本機能と利用イメージ kintone入力補助 画面イメージ

- 入力補助機能の画面イメージを以下に示す。
- 既存kintone画面にLLMへ送信ボタンを設置し、ポップアップ画面に表示された出力結果を職員が確認して、編集や転記判断が可能である。



## 4. 入力補助機能の環境構築および検証

### 4.4. 基本機能と利用イメージ プロンプト管理アプリ 画面イメージ

- 各入力補助項目に対応したLLMプロンプト（指示）を、kintone上の「プロンプト管理アプリ」で一元管理し、処理内容に応じた柔軟な支援が可能な構成とした。

プロンプト管理アプリ

レコード番号↓	ボタンのフィールドコード	入力元アプリID	入力元アプリコード	出力先アプリID	出力先アプリコード	処理タイプ	有効フラグ	RAG	Document	システムプロンプト1
120	azureButtonSpace_1_32	001	XXX	101	YYY	1	ON			# 指示 あなたは、調査員が中小企業に行ったヒアリング記録に対して
119	azureButtonSpace_6_6	001	XXX	101	YYY	6	ON	1	["下請代金支払遅延等防止法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
118	azureButtonSpace_6_5	001	XXX	101	YYY	6	ON	1	["下請代金支払遅延等防止法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
117	azureButtonSpace_6_4	001	XXX	101	YYY	6	ON	1	["下請代金支払遅延等防止法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
116	azureButtonSpace_8_8	001	XXX	101	YYY	8	ON	1	["下請中小企業振興法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
115	azureButtonSpace_7_2	001	XXX	101	YYY	7	ON	1	["下請代金支払遅延等防止法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
114	azureButtonSpace_6_3	001	XXX	101	YYY	6	ON	1	["下請中小企業振興法_utf-8.htm...]	# 指示 あなたは、調査員が中小企業に大企業との取引に関連する問題
113	azureButtonSpace_6_2	001	XXX	101	YYY	6	ON	1	["下請中小企業振興法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
112	azureButtonSpace_8_1	001	XXX	101	YYY	8	ON	1	["下請中小企業振興法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
111	azureButtonSpace_6_1	001	XXX	101	YYY	6	ON	1	["下請中小企業振興法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは
110	azureButtonSpace_7_1	001	XXX	101	YYY	7	ON	1	["下請代金支払遅延等防止法_utf-8.htm...]	以下はSYSTEMプロンプトの指示文です。[SYSTEM] # 指示 あなたは

入力補助対象のフィールドに紐づくボタン単位で管理

当該支援対象フィールドにおけるプロンプト

アプリ/フィールド毎で校閲や分類等のタスク（処理タイプ）を指定

法令等を参照する機能付与はフィールド毎に設定可能

## 4. 入力補助機能の環境構築および検証

### 4.5. 検証方法 検証手順

- 検証担当者がユースケースA～Dに該当する機能を利用し、評価結果をアンケートに記載する。

#### ① 検証の実施

- 検証担当者が、ユースケースA～Dに該当する機能を、アンケートに回答することを目的に利用する。

Azure 応答 (#内容)

出力先フィールド：事業概要等  
※ 保存時は中欄 (<標準化文章>) だけを抽出して転記します。上欄と下欄は参照用です (編集不可)。

<校閲前の文章>

・工業用ゴム製品やゴム押出部品を製造。化学、材料・資源、製造、自動車・輸送などの幅広い産業分野に製品を提供している。

<標準化文章>

【事業概要】  
(事業内容)  
・工業用ゴム製品やゴム押出部品を製造し、化学、材料・資源、製造、自動車・輸送などの幅広い産業分野に製品を提供している。

【主要事業】  
(抽出内容なし)

<ヒアリング観点指摘/標準化過程の記録>

<ヒアリング観点指摘>  
主要事業なし、状況感なし、その他情報なし

<標準化過程の記録>

内容を反映 (保存) 閉じる

#### ② 検証アンケートの実施

- 総合評価で、検証全体を通しての検証全体アンケートと、ユースケースA～Dそれぞれに対するユースケース別アンケートを行う。
- 個別評価で、検証を実施したレコードごとにアンケートを行う。

#	アンケート項目名	対象ヒアリング項目	観点	回答形式	必須/任意	説明概要	回答欄
<b>検証全体アンケート</b>							
1	検証全体アンケート全体	全AI支援対象ヒアリング項目	機能利用の意思	選択式	必須	今回実施したAI支援機能を引き続き利用した	
2			導入障壁	記述式	任意 (記入すべき内容がある場合)	なぜAI支援機能を引き続き利用したい、もしくはおたくAI支援機能を導入するにあたって、業務転換や業務機能、システム変更などは、課題などご感想を記入ください	
3			他業種への展開	記述式	任意 (記入すべき内容がある場合)	今回実施したAI支援機能の他に、今回実施したAI支援機能を利用できなかった業務があれば記入ください	
4							
<b>支援機能アンケート</b>							
5	機能検討	全AI支援対象ヒアリング項目	プロンプトの保存性	記述式	任意 (プロンプト管理アプリ担当者である場合)	適合ヒアリング項目の変更時に行うプロンプト修正を容易に行うことが困難なため	
6	機能A. 文章標準化機能 機能D. 文章整合性確認機能	【機能A】<ヒアリング先情報アプリ> 事業概要等 業界の事情等 所属業界団体 支社/海外等 備考 (自動車業界における商社等の介在)	AI出力の質	ヒアリング先情報アプリ		ヒアリング先法人番号	
7							
8							
9							
10							
11							

## 4. 入力補助機能の環境構築および検証

### 4.5. 検証方法 アンケート詳細

- 検証担当者を対象に、以下の項目のアンケートを実施した。

アンケート種別		アンケート観点	アンケート観点概要
総合評価アンケート	検証全体アンケート	継続利用の意思	<ul style="list-style-type: none"> <li>入力補助機能を引き続き利用したいと思うか。</li> </ul>
		導入障壁	<ul style="list-style-type: none"> <li>入力補助機能の導入への障壁があるか。</li> </ul>
		他業務への展開	<ul style="list-style-type: none"> <li>入力補助機能をヒアリング記録以外の業務に利用できるか。</li> </ul>
	ユースケース別アンケート	プロンプトの保守性	<ul style="list-style-type: none"> <li>法改正やヒアリング項目変更時に必要となるプロンプトの保守が容易か。</li> </ul>
		LLMの出力の質	<ul style="list-style-type: none"> <li>LLMの出力内容が理解しやすく正しいと思うか。</li> </ul>
		業務効率化・品質向上	<ul style="list-style-type: none"> <li>入力補助機能により、業務の効率化や品質向上が実現したと思うか。</li> </ul>
		UI・UX	<ul style="list-style-type: none"> <li>入力補助機能が操作しやすいと思うか。</li> </ul>
個別評価アンケート	レコード別アンケート	機能価値向上	<ul style="list-style-type: none"> <li>入力補助機能に利用価値補向上する改善点があるか。</li> </ul>
		LLM回答の妥当性	<ul style="list-style-type: none"> <li>LLMの出力内容が妥当であるか。</li> </ul>
		体感応答速度	<ul style="list-style-type: none"> <li>LLMの出力にかかる時間が許容できる範囲内か。</li> </ul>

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 検証全体アンケート

- 7割以上の検証者が入力補助機能の継続利用を望んでいる。
- しかし、LLMの出力精度、システム面の安定性、運用方法などの観点での課題もある。

アンケート概要	選択式アンケート結果	記述式アンケート結果										
<b>【継続利用の意思】</b> 入力補助機能を引き続き利用したいと思うか。	<table border="1"><thead><tr><th>回答</th><th>件数</th></tr></thead><tbody><tr><td>1: とてもそう思う</td><td>1</td></tr><tr><td>2: そう思う</td><td>15</td></tr><tr><td>3: そう思わない</td><td>3</td></tr><tr><td>4: 全くそう思わない</td><td>3</td></tr></tbody></table>	回答	件数	1: とてもそう思う	1	2: そう思う	15	3: そう思わない	3	4: 全くそう思わない	3	<ul style="list-style-type: none"><li>■ 評価された点<ul style="list-style-type: none"><li>• 文章標準化は職員等によるチェックに有効である</li><li>• カテゴリ候補は第三者の意見として参考になり有効である</li><li>• データ精度向上や業務効率化につながる。</li></ul></li><li>■ 改善が期待される点<ul style="list-style-type: none"><li>• 調査員がLLMに頼りすぎ、スキルアップにつながらない懸念がある。</li><li>• LLMが出力するまでにかかる時間が長い。</li><li>• 文章の校正はできているが、校閲までは行われておらず不十分である。</li><li>• カテゴリ候補の精度が不十分である。</li><li>• カテゴリ候補の結果を確認する必要があり、業務効率化にはつながらない。</li><li>• システムエラーが発生する場合がある。</li><li>• 参照するデータが100%正しいという前提でAIが参照してしまっている。</li></ul></li></ul>
回答	件数											
1: とてもそう思う	1											
2: そう思う	15											
3: そう思わない	3											
4: 全くそう思わない	3											
<b>【導入障壁】</b> 入力補助機能の導入への障壁があるか。	(選択式アンケートなし)	(改善点がある場合のみ回答) <ul style="list-style-type: none"><li>■ 改善が期待される点<ul style="list-style-type: none"><li>• 調査員がLLMに頼りすぎ、スキルアップにつながらない懸念がある。</li><li>• 導入には十分な精度を満たすことが必要。</li></ul></li></ul>										
<b>【他業務への展開】</b> 入力補助機能をヒアリング記録以外の業務に利用できるか。	(選択式アンケートなし)	<ul style="list-style-type: none"><li>■ 他業務への展開案<ul style="list-style-type: none"><li>• 調査対象業種、ヒアリング先の選定</li><li>• 価格交渉のフォローアップ作業</li><li>• 調査員が入力した全データの集計・分析作業でのAI支援</li></ul></li></ul>										

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースA）

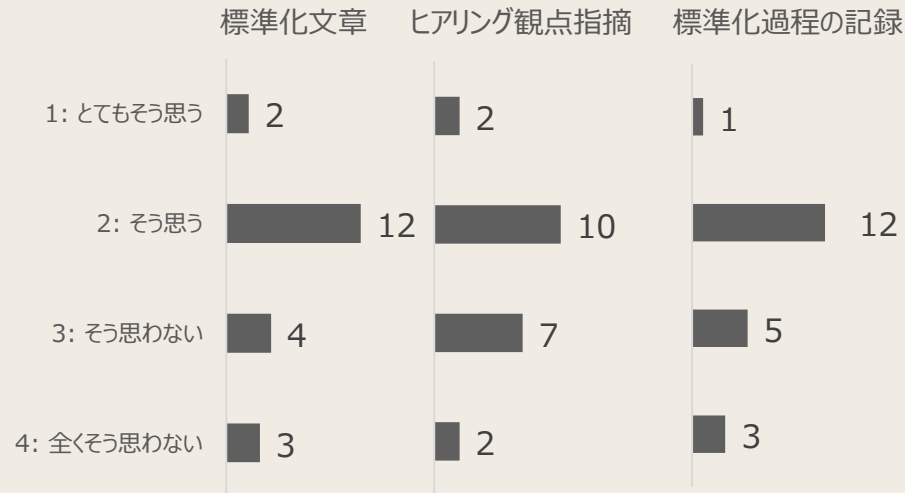
- 出力の質については、標準化の水準に関する課題がある。
- 業務効率化、品質向上については、一定の効果が見込まれるという評価もあるが、処理時間、運用方法に関する課題がある。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

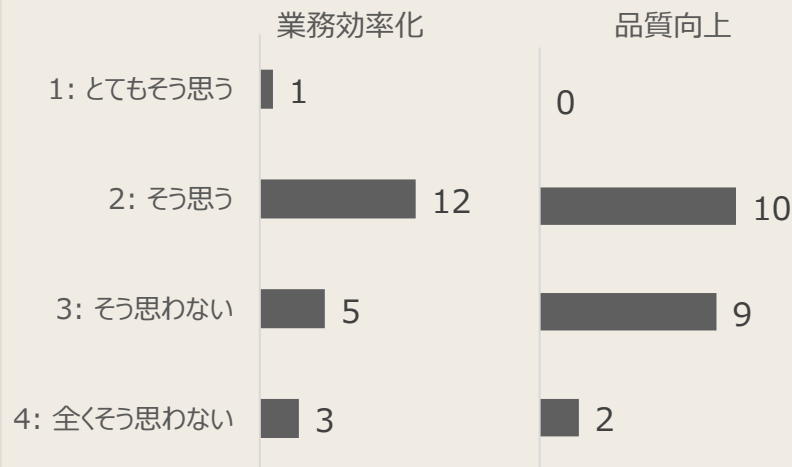
**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
  - ・ ヒアリング観点の指摘が不十分である。
  - ・ 文章の校正はできているが、校閲までは行われておらず不十分である。
  - ・ すでに構造化されている文章であると、逆に読みにくくなる場合がある。

**【業務効率化・品質向上】**  
入力補助機能の導入への障壁があるか。



■ 評価された点

- ・ 文章の体裁の標準化で品質向上が見込まれる。
- ・ LLMが文章を標準化することで入力作業時間が短縮される。
- ・ 文章の誤記が減少することでチェックの工数が削減される。

■ 改善が期待される点

- ・ 文章の体裁レベルの修正で、品質向上・業務効率化は期待できない。
- ・ LLM頼りすぎた入力に陥る可能性がある。
- ・ LLMの処理時間が長く、業務効率化にならない。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースA）

- UI・UXについては、概ね問題ないという評価だが、出力結果の表示方法に関する課題もある。

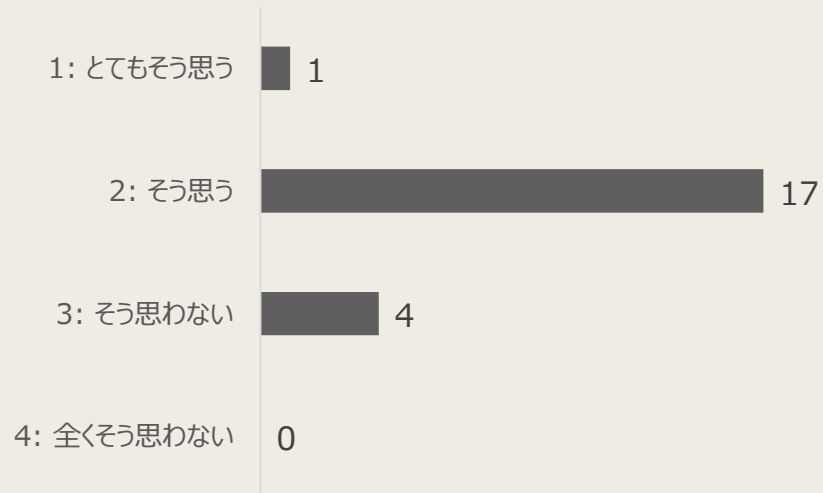
#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

##### 【UI・UX】

入力補助機能が操作しやすいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
- 結果が長い場合、全体を一括で見られない。

##### 【機能価値向上】

入力補助機能に利用価値補向上する改善点があるか。

(選択式アンケートなし)

(改善点がある場合のみ回答)

- 改善が期待される点
- 入力文章と標準化文章の配置や表示方法を工夫してほしい。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースA）

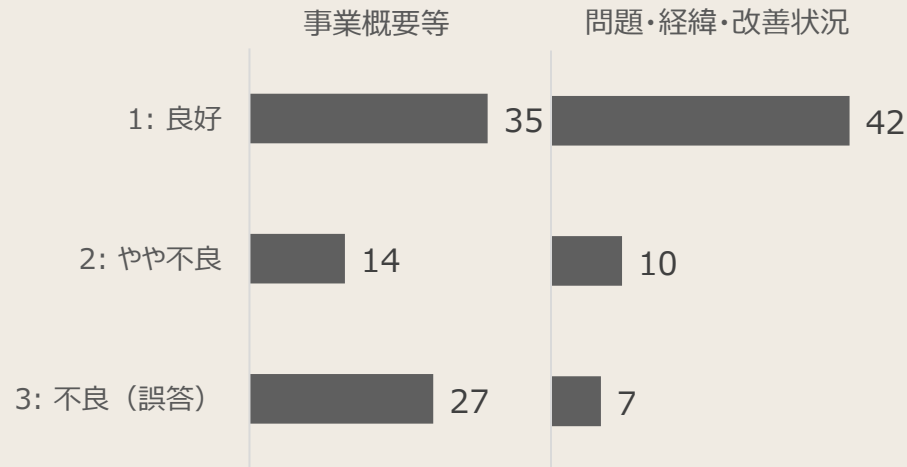
- システム面の安定性の問題は特定のヒアリング項目で生じていると考えられる。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

【LLM回答の妥当性】  
LLMの出力内容が妥当であるか。

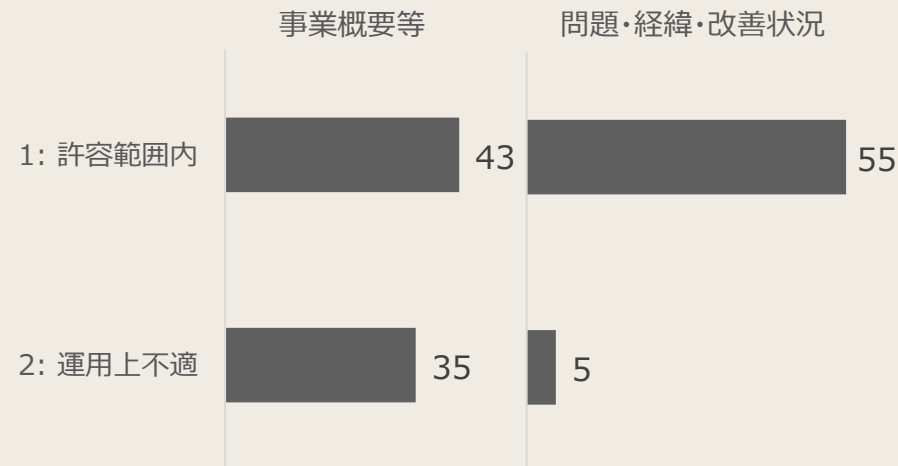


（改善点がある場合のみ回答）

■ 改善が期待される点

- 文章量が多い場合にLLMの出力が表示されないエラーが発生する。
- 文章の校正はできているが、校閲は行われていない。
- 専門用語に対応できていない場合がある。

【体感応答速度】  
LLMの出力にかかる時間が許容できる範囲内か。



（改善点がある場合のみ回答）

■ 改善が期待される点

- LLMが出力するまでにかかる時間が長い。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースB）

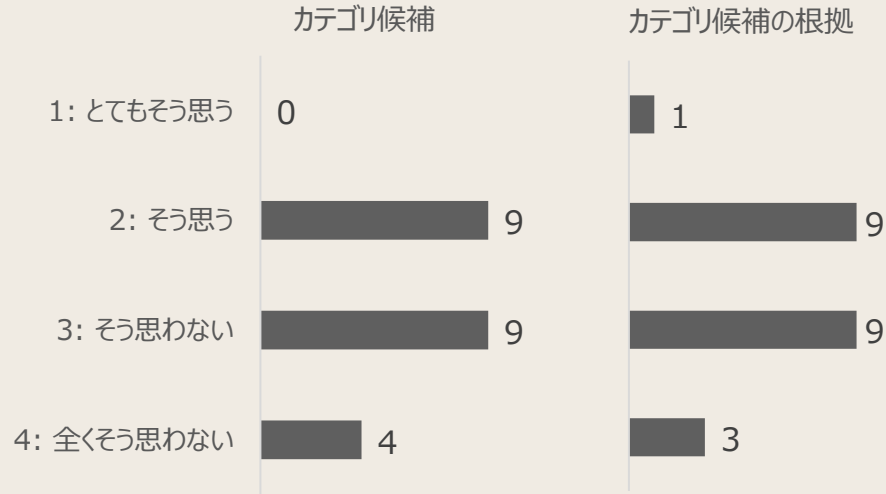
- 出力の質については、複数候補の優劣明示や業務固有ルールへの適応不足に関する課題がある。
- 業務効率化、品質向上については、一定の効果が見込まれるという評価もあるが、出力の質や処理時間、運用方法に関する課題がある。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。

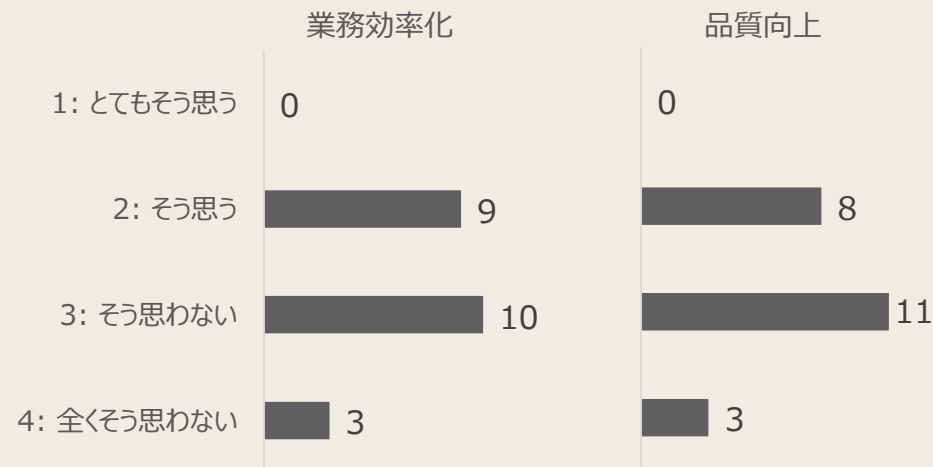


(改善点がある場合のみ回答)

■ 改善が期待される点

- 業種判定ルールに対応できていない場合がある。
- 参照する文章の全体からではなく、一部のみから業種を判定している場合がある。
- 正しい選択肢が1つ目の候補として出力されない場合がある。
- 参照するデータの情報が少ない場合、憶測で推論している。

**【業務効率化・品質向上】**  
入力補助機能の導入への障壁があるか。



■ 評価された点

- 調査員自身が選択した候補が正しいかを確認することで、品質向上は可能と考えられる。
- 候補提示により、チェックの時間が短縮される。

■ 改善が期待される点

- 調査員が確認を行わないと、LLMが誤った出力をした場合に品質が落ちる可能性がある。
- 適切でない候補も含めて、候補が複数提示されるため、業務が効率化しない。
- 推論時間が長いため、業務が効率化しない。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースB）

- UI・UXについては、概ね問題ないという評価だが、出力結果の表示方法に関する課題もある。

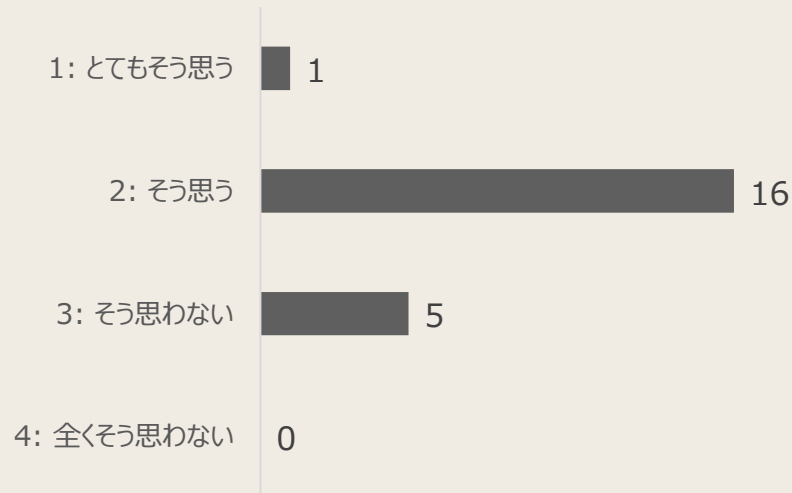
#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

##### 【UI・UX】

入力補助機能が操作しやすいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
- 候補が複数提示されていることがわかりにくい。
- 業種分類は分類コードも表示されないとわかりにくい。

##### 【機能価値向上】

入力補助機能に利用価値補向上する改善点があるか。

(選択式アンケートなし)

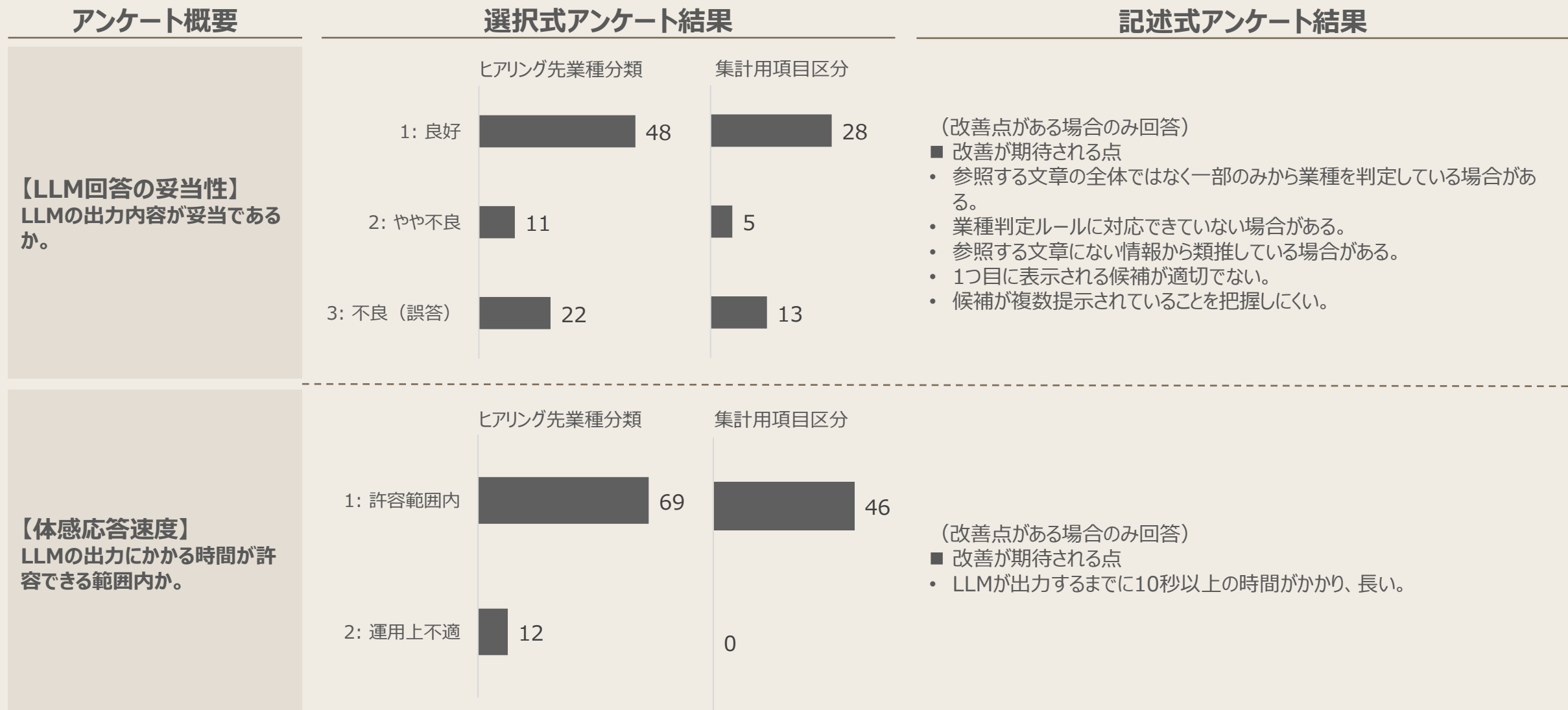
(改善点がある場合のみ回答)

- 改善が期待される点
- 複数候補の優劣の情報が欲しい。
- 複数の候補と根拠を一括で見たい。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースB）

- 業務固有ルールへの適応不足やシステム面の安定性の問題は特定のヒアリング項目で生じていると考えられる。



## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースC）

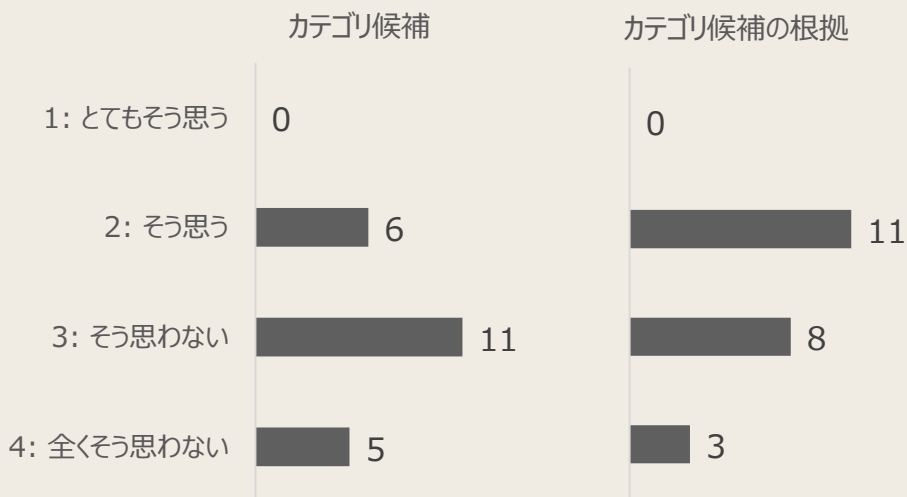
- 出力の質については、複数候補の優劣明示や業務固有ルールへの適応不足に関する課題がある。
- 業務効率化、品質向上については、一定の効果が見込まれるという評価もあるが、出力の質や処理時間、運用方法に関する課題がある。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

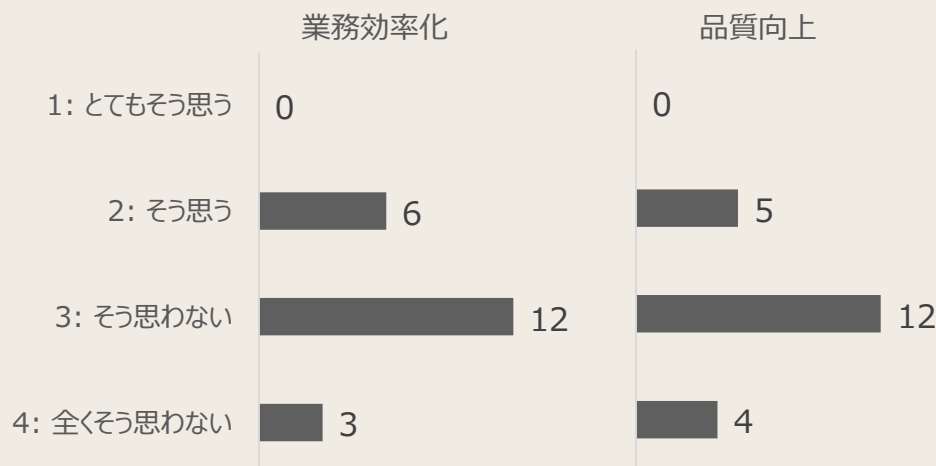
**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
  - ・ ヒアリング結果に対する所見に誤りがある。
  - ・ 知的財産関連の所見に誤りが多い。
  - ・ 資本金区分や自家使用、NBを考慮できていない。
  - ・ 参照するデータの情報量が少ない場合、憶測で推論している。

**【業務効率化・品質向上】**  
入力補助機能の導入への障壁があるか。



- 評価された点
  - ・ 調査員自身が選択した候補が正しいかを確認することで品質向上は可能と考えられる。
  - ・ 十分な精度での候補が得られる場合は、チェックの時間が短縮される。

- 改善が期待される点
  - ・ LLMのカテゴリ候補と調査員の所見が異なる場合、判断に時間がかかる。
  - ・ 調査員が確認を行わないと、LLMが誤った出力をした場合に品質が落ちる可能性がある。
  - ・ 適切でない候補も含めて、候補が複数提示されるため、業務が効率化しない。
  - ・ 推論時間が長い場合、業務が効率化しない。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユーケース別アンケート（ユースケースC）

- UI・UXについては、概ね問題ないという評価だが、出力結果の表示方法に関する課題もある。

アンケート概要	選択式アンケート結果	記述式アンケート結果										
<b>【UI・UX】</b> 入力補助機能が操作しやすいと思うか。	<table border="1"><thead><tr><th>回答</th><th>件数</th></tr></thead><tbody><tr><td>1: とてもそう思う</td><td>0</td></tr><tr><td>2: そう思う</td><td>14</td></tr><tr><td>3: そう思わない</td><td>6</td></tr><tr><td>4: 全くそう思わない</td><td>1</td></tr></tbody></table>	回答	件数	1: とてもそう思う	0	2: そう思う	14	3: そう思わない	6	4: 全くそう思わない	1	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• 候補が複数提示されていることがわかりにくい。</li><li>• 関連する複数項目が独立して推論されると、操作上の不整合が起こる。</li><li>• 複数項目を同時に推論すると、推論結果が確認しづらい。</li><li>• 上位の「ヒアリング先情報アプリ」「取引先情報アプリ」で更新した内容が反映されていない。</li></ul>
回答	件数											
1: とてもそう思う	0											
2: そう思う	14											
3: そう思わない	6											
4: 全くそう思わない	1											
<b>【機能価値向上】</b> 入力補助機能に利用価値補向上する改善点があるか。	(選択式アンケートなし)	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• 資本金区分や自家使用、NBを考慮してほしい。</li><li>• 複数候補の優劣の情報が欲しい。</li></ul>										

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースC）

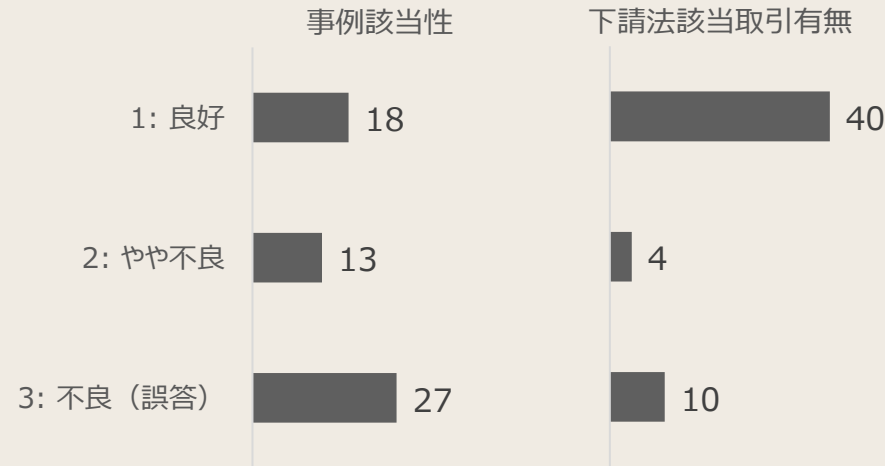
- 業務固有ルールへの適応不足やシステム面の安定性の問題はユースケースCに対応する複数のヒアリング項目で生じていると考えられる。

#### アンケート概要

#### 選択式アンケート結果

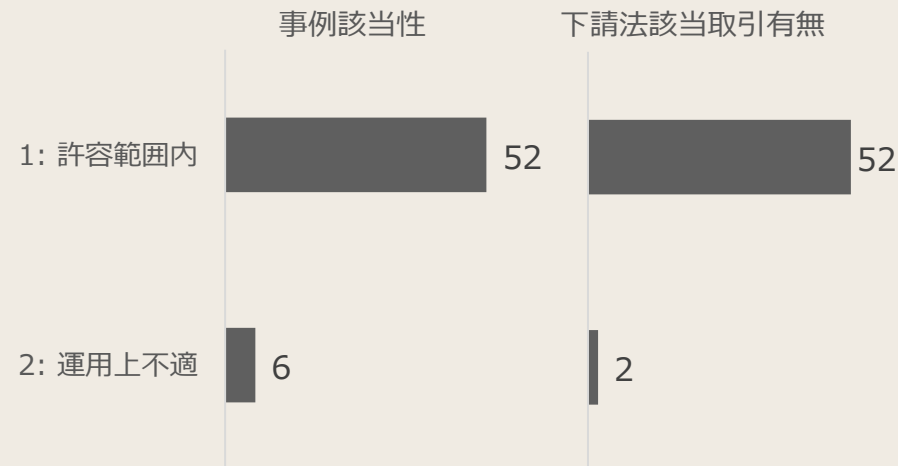
#### 記述式アンケート結果

【LLM回答の妥当性】  
LLMの出力内容が妥当であるか。



- (改善点がある場合のみ回答)
- 改善が期待される点
  - ・ 所見が誤っている場合がある。
  - ・ 1つ目を選択されている候補が適切でない。
  - ・ 候補が複数提示されていることを把握しにくい。
  - ・ 資本金区分や自家使用、NBを考慮できていない。

【体感応答速度】  
LLMの出力にかかる時間が許容できる範囲内か。



- (改善点がある場合のみ回答)
- 改善が期待される点
  - ・ LLMが出力するまでにかかる時間が長い。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースD）

- 出力の質については、参照データ依存の出力品質に関する課題がある。
- 業務効率化、品質向上については、一定の効果が見込まれるという評価もあるが、出力の質に関する課題がある。

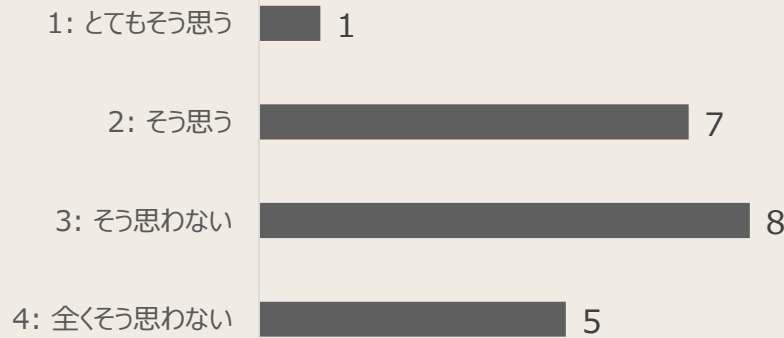
#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。

整合性確認結果



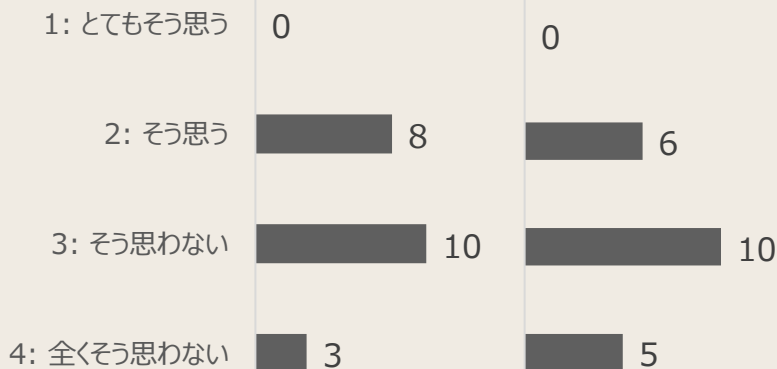
(改善点がある場合のみ回答)

- 改善が期待される点
  - 参照するデータを100%正として判断しているため、参照先の文章に誤りが含まれない必要がある。
  - 文章量が多い場合に、LLMの出力が表示されないエラーが発生する。

**【業務効率化・品質向上】**  
入力補助機能の導入への障壁があるか。

業務効率化

品質向上



■ 評価された点

- 整合性チェックにより、記入した文章の誤りに気付くことで、品質が向上する。

■ 改善が期待される点

- 参照するデータが誤っている場合、誤った判断を招き、品質が低下する可能性がある。
- 整合性確認結果を確認する手間がかかり、業務効率が低下する。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースD）

- UI・UXについては、概ね問題ないという評価だが、出力結果の表示方法に関する課題もある。

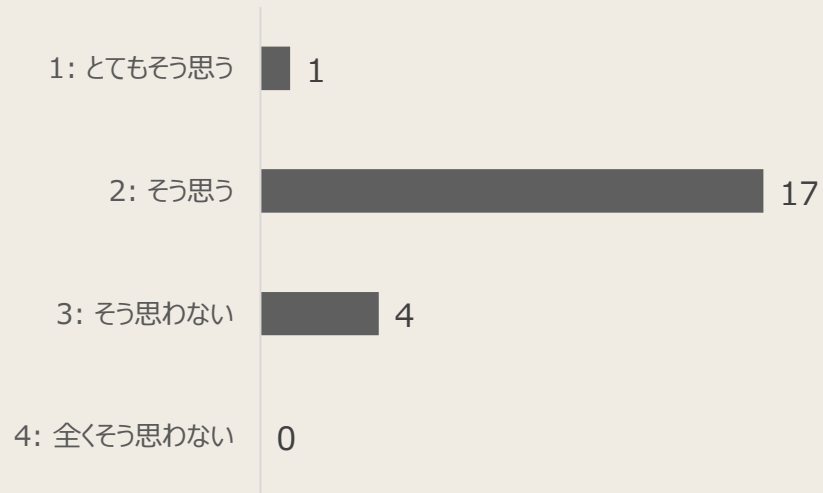
#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

##### 【UI・UX】

入力補助機能が操作しやすいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
- 結果が長い場合、全体を一括で見られない。

##### 【機能価値向上】

入力補助機能に利用価値補向上する改善点があるか。

(選択式アンケートなし)

(改善点がある場合のみ回答)

- 改善が期待される点
- もう少し簡潔に結果を表示してほしい。

## 4. 入力補助機能の環境構築および検証

### 4.6. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースD）

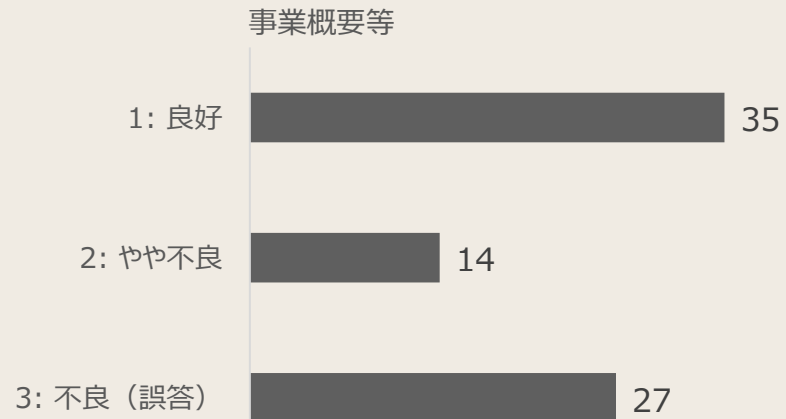
- 参照データ依存の出力品質の問題は、ユースケースDに対応するヒアリング項目全体で生じていると考えられる。

#### アンケート概要

【LLM回答の妥当性】  
LLMの出力内容が妥当であるか。

【体感応答速度】  
LLMの出力にかかる時間が許容できる範囲内か。

#### 選択式アンケート結果



#### 記述式アンケート結果

(改善点がある場合のみ回答)

■ 改善が期待される点

- 参照するデータを100%正として判断しているため、参照先の文章に誤りが含まれない必要がある。
- 文章量が多い場合に、LLMの出力が表示されないエラーが発生する。

(改善点がある場合のみ回答)

■ 改善が期待される点

- LLMが出力するまでにかかる時間が長い。

## 4. 入力補助機能の環境構築および検証

### 4.7. 今後対応すべき課題

- アンケート結果から品質・UX・運用に関する7つの課題を洗い出した。
- 各課題に対して対応策を検討する必要がある。

課題名	課題名	課題概要	対応アンケートコメント例	ユースケース
意思決定支援としての品質	入力元情報の品質向上	<ul style="list-style-type: none"> <li>LLMの入力元となるデータの質または量が十分でなく、分類などの出力結果の精度が上がらない。</li> </ul>	<ul style="list-style-type: none"> <li>参照するデータを100%正として判断しているため、参照先の文章に誤りが含まれない必要がある。</li> <li>参照するデータの情報量が少ない場合、憶測で推論している。</li> </ul>	B / C / D
	優先順位付け	<ul style="list-style-type: none"> <li>複数の候補すべてを確認する必要があり、その確認工数が大きい。</li> </ul>	<ul style="list-style-type: none"> <li>適切でない候補も含んだ候補が複数提示されるため、業務が効率化しない。</li> </ul>	B / C
	文章標準化の品質向上	<ul style="list-style-type: none"> <li>LLMで生成した標準化文章が、業務に適用可能なレベルに至っていない。</li> </ul>	<ul style="list-style-type: none"> <li>文章の校正はできているが、校閲までは行われておらず不十分である。</li> <li>すでに構造化されている文章であると、逆に読みにくくなる場合がある。</li> <li>ヒアリング観点の指摘が不十分である。</li> </ul>	A
	専門知識の考慮	<ul style="list-style-type: none"> <li>専門用語を含む文章の校閲において、専門的な知識が必要となる処理を適切に行えていない</li> </ul>	<ul style="list-style-type: none"> <li>業種判定ルールに沿っていない場合がある。</li> <li>専門用語に対応できていない場合がある。</li> </ul>	A
	法令基準の正確な反映	<ul style="list-style-type: none"> <li>法令で定められた基準に基づくカテゴリ判定が十分に行えていない。</li> </ul>	<ul style="list-style-type: none"> <li>知的財産権や資本金区分、自家使用、NBを考慮した所見ではない。</li> </ul>	B / C
システムにおけるUX	応答時間の短縮	<ul style="list-style-type: none"> <li>LLMの処理時間が長く、業務効率化および円滑な利用を妨げている。</li> </ul>	<ul style="list-style-type: none"> <li>LLMの処理時間が長く、業務効率化にならない。</li> </ul>	A / B / C / D
	システムのエラー耐性向上	<ul style="list-style-type: none"> <li>機能が正常に動作せず、想定した利用シナリオで機能を利用できないケースがある。</li> </ul>	<ul style="list-style-type: none"> <li>文章量が多い場合に、LLMの出力が表示されないエラーが発生する。</li> <li>上位の「ヒアリング先情報アプリ」「取引先情報アプリ」で更新した内容が下位の「テーマ別ヒアリングアプリ」に反映されていない。</li> </ul>	A / C / D
運用・ガバナンス（人材育成）	専門性・判断力のスキルアップ	<ul style="list-style-type: none"> <li>LLMへの過度な依存は、職員の専門性低下やヒアリング記録の品質低下を引き起こすリスクがある。</li> </ul>	<ul style="list-style-type: none"> <li>調査員が頼りすぎ、スキルアップにつながらない懸念がある。</li> <li>調査員が確認を行わないと、LLMが誤った出力をした場合に品質が落ちる可能性がある。</li> </ul>	A / B / C / D

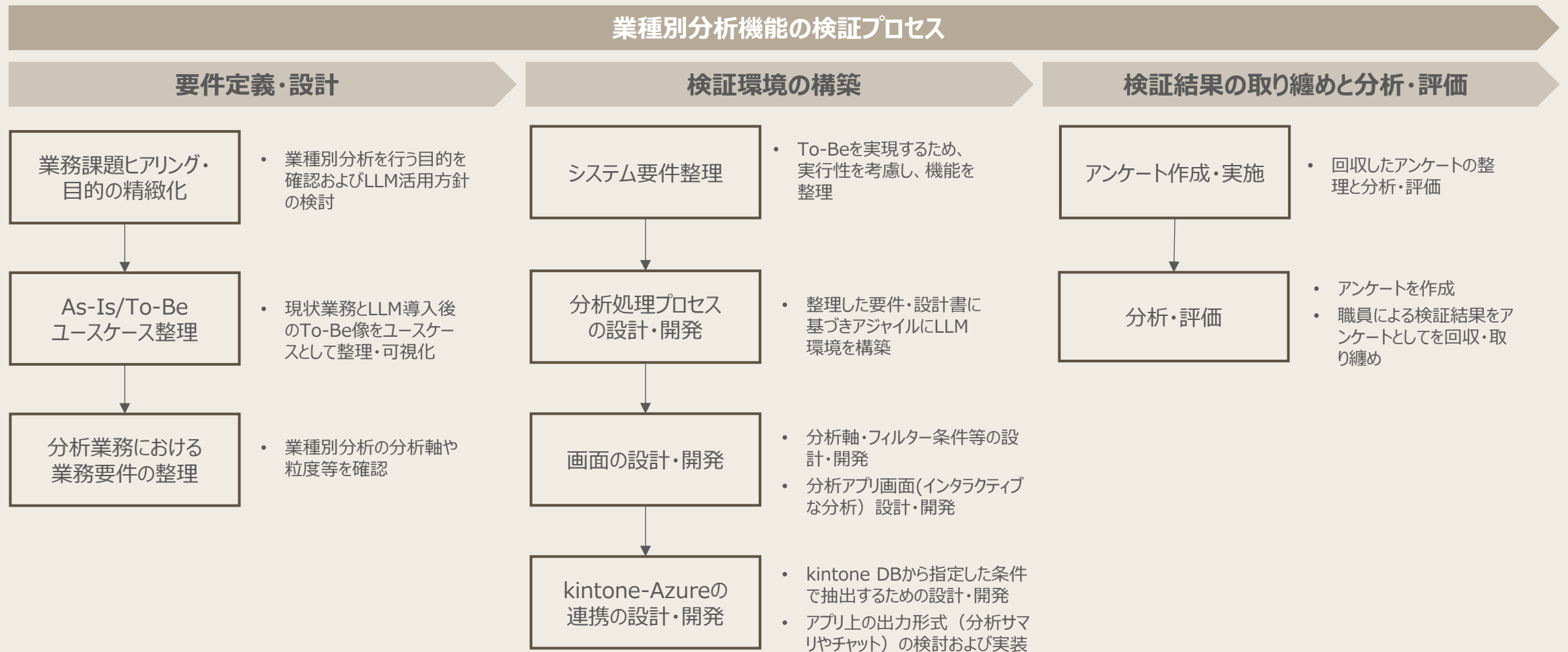
# 5. 業種別分析機能の環境構築および検証

- 5.1. 検証の実施方針
- 5.2. 検証対象ユースケースの概要
- 5.3. 検証環境の要件
- 5.4. 検証環境の構成・仕様
- 5.5. 基本機能と利用イメージ
- 5.6. 検証方法
- 5.7. 検証結果
- 5.8. 今後対応すべき課題

## 5. 業種別分析機能の環境構築および検証

### 5.1. 検証の実施方針

- 分析業務に関する要件定義を実施し、検証環境を構築した。本検証は、kintoneに蓄積されたヒアリングデータを活用した機能を実装している。業務有用性については、アンケートを設計・実施し、その結果に基づき評価を行う。



## 5. 業種別分析機能の環境構築および検証

### 5.2. 検証対象ユースケースの概要 検証対象ユースケースの整理 (1/2)







- 業種別分析に関する8つのユースケースのうち、効果と実現可能性が高いと見込まれる「分析軸に沿ったヒアリング結果の抽出・要約・示唆提示」「条件一致ヒアリング結果抽出」を検証対象とした。

ユースケース名	ユースケース概要	効果	実現可能性	実施優先度
I 分析軸に沿ったヒアリング結果の抽出・要約・示唆提示	<ul style="list-style-type: none"> <li>主なヒアリング結果の抽出や要約、ガイドライン作成・啓発などに向けた示唆出し</li> </ul>	<ul style="list-style-type: none"> <li>資料作成の工数削減・品質向上</li> </ul>	<ul style="list-style-type: none"> <li>設計・実装可能</li> </ul>	高
II 条件一致ヒアリング結果抽出	<ul style="list-style-type: none"> <li>洗い出したい特徴や条件をチャット形式で入力し、合致するヒアリング結果をすべて抽出</li> </ul>	<ul style="list-style-type: none"> <li>頻発する調査の工数削減</li> <li>暗黙知の可視化</li> </ul>	<ul style="list-style-type: none"> <li>資産を流用した形で実装可能</li> </ul>	高
企業への助言生成	<ul style="list-style-type: none"> <li>好事例の傾向から、取引適正化に向けた取組の示唆出し</li> </ul>	<ul style="list-style-type: none"> <li>信頼性の高い示唆により企業のキャッシュフローが改善</li> </ul>	<ul style="list-style-type: none"> <li>検証期間が不十分</li> </ul>	低
ヒアリング結果分析とアンケート結果の新規項目作成	<ul style="list-style-type: none"> <li>ヒアリング結果分析から、アンケートの新規項目を提案</li> <li>アンケート結果から、深掘りするヒアリング項目を提案</li> </ul>	<ul style="list-style-type: none"> <li>アンケートとヒアリングの相互的なブラッシュアップ</li> </ul>	<ul style="list-style-type: none"> <li>追加でアンケートの情報が必要</li> <li>新規項目への検証が必須</li> </ul>	低
ヒアリング結果の属性別傾向分析・示唆出し	<ul style="list-style-type: none"> <li>業種、企業規模、地域などを軸にしたヒアリング結果の傾向分析</li> <li>傾向分析からガイドライン作成・啓発などに向けた示唆出し</li> </ul>	<ul style="list-style-type: none"> <li>表面的な集計・抽出では見えない傾向や示唆を提示</li> </ul>	<ul style="list-style-type: none"> <li>分析結果・示唆に対して十分な検証が必要</li> </ul>	中

## 5. 業種別分析機能の環境構築および検証

### 5.2. 検証対象ユースケースの概要 検証対象ユースケースの整理 (2/2)

- 業種別分析に関する8つのユースケースのうち、効果と実現可能性が高いと見込まれる「分析軸に沿ったヒアリング結果の抽出・要約・示唆提示」「条件一致ヒアリング結果抽出」を検証対象とした。

ユースケース名	ユースケース概要	効果	実現可能性	実施優先度
ヒアリング結果分析から新規ヒアリング項目作成	<ul style="list-style-type: none"><li>・ ヒアリング結果分析から、深掘りすべきヒアリング項目を提案</li></ul>	<ul style="list-style-type: none"><li>・ ヒアリング項目のブラッシュアップ </li></ul>	<ul style="list-style-type: none"><li>・ 新規項目への検証が必須 </li></ul>	低
アンケート・ヒアリング結果分析から示唆出し	<ul style="list-style-type: none"><li>・ アンケートとヒアリング結果の双方から出た示唆を組み合わせたガイドライン作成・啓発などに向けた示唆出し</li></ul>	<ul style="list-style-type: none"><li>・ アンケートとヒアリングの示唆の共通点と違いを可視化</li><li>・ 隠れた示唆を提示 </li></ul>	<ul style="list-style-type: none"><li>・ 追加でアンケートの情報が必要 </li></ul>	中
調査員のヒアリング記録誤り分析	<ul style="list-style-type: none"><li>・ 調査員がどのような項目で記録ミスをしやすいかの分析と、教育、マニュアル改訂に向けた示唆出し</li></ul>	<ul style="list-style-type: none"><li>・ 記録ミスの低減（入力補助機能により今後の利用価値が下がる） </li></ul>	<ul style="list-style-type: none"><li>・ 設計・実装可能 </li></ul>	低

## 5. 業種別分析機能の環境構築および検証

### 5.2. 検証対象ユースケースの概要 ユースケース I .分析軸に沿ったヒアリング結果の抽出・要約・示唆提示

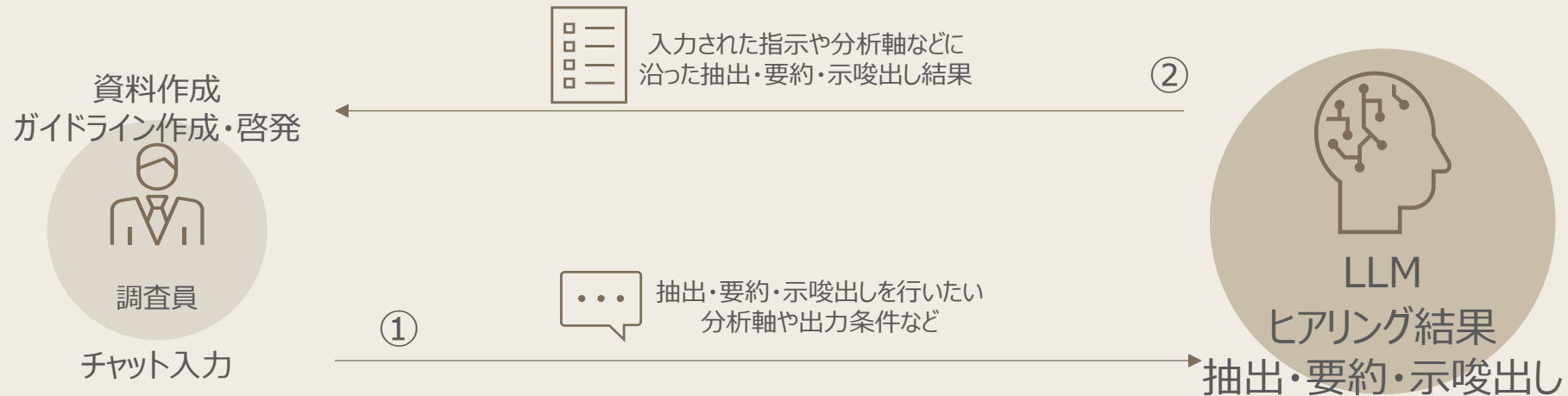
- 調査員がチャットで送った処理の指示や条件に従って、LLMがヒアリング結果の抽出や要約、示唆出しを行う。
- これにより資料作成やガイドライン作成等に要する工数の削減、品質向上を実現する。

#### Before

- 大量のヒアリング結果の中から分析軸に沿ったアリング結果を抽出し、要約、示唆出しを行う作業に工数がかかっている。
- ヒアリング結果が多く、職員によって抽出するヒアリング結果がばらつく可能性がある。

#### After

- 自動抽出されたヒアリング結果を利用することで、資料作成やガイドライン作成等に要する工数の削減や品質向上が実現する。



## 5. 業種別分析機能の環境構築および検証

### 5.2. 検証対象ユースケースの概要 ユースケースⅡ.条件一致ヒアリング結果抽出

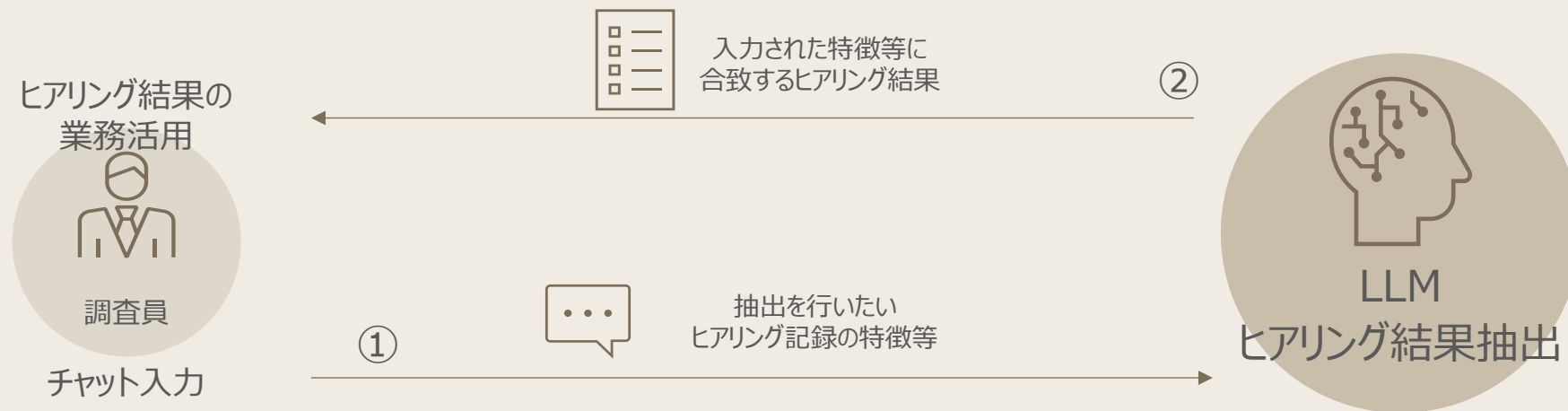
- 調査員がチャットで送った抽出条件に従って、LLMが代表的なヒアリング結果を抽出する。
- これにより調査員のヒアリング結果の抽出に要する工数が削減される。

#### Before

- 大量のヒアリング結果の中から参照したいヒアリング結果を探す作業に工数がかかっている。
- ヒアリング結果をキーワード以外で探せず、事例の抽出が不十分になっている。

#### After

- ヒアリング結果の特徴や条件の入力のみで、参照したいヒアリング結果の抽出が完了する。



## 5. 業種別分析機能の環境構築および検証

### 5.3. 検証環境の要件

- ヒアリング結果の効率的な抽出・要約・示唆出しを実現するための主要機能要件を整理した。

機能要件		詳細	実現方法
コンテンツ生成機能	プロンプト生成機能	プロンプトの知見有無に関係なく、LLMに適切な処理を行わせるためのプロンプトを簡易的に作成できること。また、自由にプロンプトを編集できること。	Excelマクロを用いたプロンプト作成ツールを作成し、職員による円滑な検証を実現する。
	抽出・要約・示唆提示機能	プロンプト指示に基づき、ヒアリング結果のデータを参照し、抽出結果・要約・示唆を生成し、ユーザーに出力できること。 ※ユースケースII: ヒアリング結果洗い出しチャットは、 <u>抽出結果の生成のみ</u>	kintoneから所望の条件でのレコード抽出機能を具備した業種別分析を実行するためのWebアプリケーションを構築する。
	対象絞り込み機能	調査日・業種・法人番号・資本金・従業員数・ヒアリング先中小企業・事例区分でDBの参照範囲を指定できること。（参照する調査日は5年前までとする。）	整理した要件に準じてレコード抽出機能を実装する。
	結果推敲機能	対話型のインターフェースにより、LLMからの回答を継続的に推敲できること。	出力された分析結果に対して結果を推敲可能な対話型のインターフェース機能を実装する。
その他機能	ログイン機能	本検証フェーズでは未実施	—
	ログ蓄積機能	本検証フェーズでは未実施	—
	利用状況照会機能	本検証フェーズでは未実施	—

## 5. 業種別分析機能の環境構築および検証

### 5.4. 検証環境の構成・仕様 システム構成図

- Webアプリ画面より、過去ヒアリングデータなどの分析を支援するための処理やLLMを組み込み、検証環境を準備した。

#### 業種別分析機能



## 5. 業種別分析機能の環境構築および検証

### 5.5. 基本機能と利用イメージ 業種別分析機能イメージ

- 業種別分析機能は、プロンプトを作成するExcelマクロと、LLMを用いてkintoneアプリのDBから抽出した結果に対する分析をするアプリで構成した。

プロンプト作成マクロ



業種別分析機能アプリ



**実行**

**機能設定 (手順①)**  
 抽出  要約  示唆

**処理条件設定 (手順②)**  
抽出件数 **プロンプト**

**事例区分設定**  
 好事例 (選択)

**分析軸設定**  
 取引先  
 支払

**# 指示**  
あなたは、調査員が入力した中小企業に対する取引ヒアリング記録から、情報を抽出・要約するアシスタントです。

**## 抽出**  
以下の分析軸に関連する代表的な事例をそれぞれ 20 件抽出してください。また、抽出した事例のユニークキーも付加してください。

**## 要約**  
以下の分析軸に関連する事例をそれぞれ 200 文字で要約してください。また、すべての分析軸を通しての要約を 200 文字で行ってください。

**## 分析軸**  
- 取引先上位との関係の価格転嫁に関連する好事例

**# 入力データ形式**  
ヒアリング記録文章です。

**中企庁取引Gメン分析機能 (検証事業)**

**注意事項**

① 取扱禁止の情報  
・ 機密性 3 情報  
・ 生成AI規約で禁止されている著作物  
・ 登録中備ほか倫理的に問題のある内容

② 出力結果の取扱い  
・ 生成AIの出力結果は職員が確認せず、そのまま資料化・行政文書化することは禁止

メッセージを入力して [Enter] で送信、[Shift] + [Enter] で改行します。

+ 中企庁 業種別分析機能

● 生成AIの出力結果には不正確な情報が含まれる場合があります。必ず職員が確認してください。

現在の会話セッションID: 11697093145841

## 5. 業種別分析機能の環境構築および検証

### 5.5. 基本機能と利用イメージ 業種別分析機能イメージ

- 分析業務における分析軸や条件を協議した結果、多様なプロンプトに対応する必要があることがわかった。
- 本検証においては、職員の意図に沿ったプロンプトの作成補助をクイックに検証するために、Excelマクロを採用した。

#### プロンプト作成マクロ

**業種分析プロンプト生成アプリ**

**実行**

**機能設定 (手順①)**

抽出  要約  示唆

**処理条件設定 (手順②)**

抽出件数:  件 要約文字数:  文字 示唆タイプ:

**事例区分設定 (手順③)**

好事例  注意を要する事例 (選択不可)

**分析軸設定 (手順④)**

取引先上位との関係

支払条件  債権転讓  債権償還  債取引等  各種取引

行政関係との取引

支払条件  債権転讓  債取引等  各種取引

支払条件

支払遅延  手数料負担  減価・引当金 (引当金)

100%の取扱い  現金化

親等の取引

修正費用負担  保有権  器具の売却  分別払い

保証・資産質借・保証仲介等

**各種取引関係**

保証委託業務  IT・在庫対応  検査・品質管理  コスト削減

法令遵守体制  取引先・協賛金・小売決済等

**知財取引関係**

秘密情報  権利の譲渡  産業対価

**プロンプト**

あなたは、調査員が入力した中小企業に対する取引ヒアリング記録から、情報を抽出・要約するアシスタントです。

## 抽出  
以下の分析軸に関連する代表的な事例をそれぞれ 20 件抽出してください。また、抽出した事例のユニークキーも付加してください。

## 要約  
以下の分析軸に関連する事例をそれぞれ 200 文字で要約してください。  
また、すべての分析軸を通しての要約を 200 文字で行ってください。

## 分析軸  
- 取引先上位との関係の債権転讓に関する好事例

## 入力データ形式  
ヒアリング記録文章です。

## 出力データ形式  
入れる値がない場合は、空欄で出力してください。

```
{
  "個別分析": [
    {
      "分析軸名": "取引先上位との関係の債権転讓に関する好事例",
      "抽出結果": [
        {
          "登録ユニークキー": "",
          "事例": "(抽出されたテキスト・内容)"
        }
      ]
    },
    {
      "要約結果": {
        "要約結果": "(分析軸の要約)"
      }
    },
    {
      "示唆結果": {
        "示唆結果": "(分析軸から導かれる示唆)"
      }
    }
  ],
  "総括": {
    "要約": "(すべての分析軸を総括した要約)",
    "示唆": "(総合的な示唆)"
  }
}
```

**実行**

**処理条件設定 (手順①)**

抽出件数:  件

**事例区分設定 (手順②)**

好事例  注意を要する事例

**抽出条件 (手順③)**

条件①:

条件②:

条件③:

条件④:

条件⑤:

**プロンプト**

# 指示  
あなたは、調査員が入力した中小企業に対する取引ヒアリング記録から、情報を抽出・要約するアシスタントです。

## 抽出  
以下の分析条件に関連する代表的な事例をそれぞれ 3 件抽出してください。また、抽出した事例のユニークキーも付加してください。

## 分析条件  
- 支払条件に関する事例(好事例)  
- 支払遅延に関する事例(好事例)

# 入力データ形式  
ヒアリング記録文章です。

# 出力データ形式  
入れる値がない場合は、空欄で出力してください。

```
{
  "分析軸出力": [
    {
      "分析軸名": "支払条件に関する事例(好事例)",
      "抽出": [
        {
          "ユニークキー": "",
          "抽出結果": "(抽出されたテキスト・内容)"
        }
      ]
    },
    {
      "要約": "(空欄)",
      "示唆": "(空欄)"
    }
  ],
  "分析軸名": "支払遅延に関する事例(好事例)",
}
```

Excel上で、機能設定・処理設定・分析軸を選択・入力

選択・入力した条件に沿ったプロンプトがExcel上に出力

## 5. 業種別分析機能の環境構築および検証

### 5.5. 基本機能と利用イメージ 業種別分析機能イメージ

- 業種別分析機能は、LLMを用いてkintoneから抽出した分析対象に対して、職員が指定した分析軸に沿った分析結果を出力する。

kintoneから抽出用の条件設定画面

**分析機能**  
(ヒアリング先企業の条件設定)

**調査日 (期間指定) (必須)**  
※ 検証事業においては、2025/5/1 ~ 2026/1/15 の範囲で選択してください。  
yyyy/mm/dd ~ yyyy/mm/dd

**業種 (大分類)**  
指定なし

**業種 (中分類)**  
指定なし

**法人番号**  
0123456789012

**資本金 (万円)**  
下限 ~ 上限

**従業員数**  
下限 ~ 上限

**ヒアリング先中小企業 (訪問先都道府県)**  
指定なし

**事例区分 (必須)**  
選択してください (必須)

キャンセル 決定

分析結果の表示画面

# 指示  
あなたは、調査員が入力した中小企業に対する取引ヒアリング記録から、情報を抽出・要約するアシスタントです。

## 抽出  
以下の分析軸に関連する代表的な事例をそれぞれ20件抽出してください。また、抽出した事例のユニークキーも付加してください。

## 要約  
以下の分析軸に関連する事例をそれぞれ200文字で要約してください。  
また、すべての分析軸を通しての要約を200文字で行ってください。

Copy All Edit

**分析軸：取引先上位との関係の価格転嫁に関する好事例**

**要約**  
支払条件の長期化、不明確な取決め、慣習化した不利な条件、保管に関する過剰な負担などが長期的に続き、改善を申し入れにくい状況が見られる。

**示唆**  
長期的な支払サイト、一方的な減額、費用負担の押し付け等が慣習として残存しやすく、組織間の力関係や依存構造が改善阻害要因となりやすい。

**抽出結果**

**事例 1**  
発注数量を指定された期限までに準備しても、相手先がその期限に受け取らないため、売上計上や請求ができず、保管のみが続くケースがある。これは長年の慣行として担当者間で行われており、改善を申し入れるのは難しい状況となっている。他の同業種でも同様の状況があると推測される。

項目	内容
調査日	20XX-XX-XX
担当職	XXX経済産業局
登録ユニークキー	XXXXXXXXXXXXXX
ヒアリング先	事業者A
取引先	事業者B

**事例 2**

メッセージを入力して [Enter] で送信、[Shift]+[Enter] で改行します。

指示プロンプトおよびkintoneの抽出等の結果をLLMに入力

分析軸や分析指示に沿って、分析結果を表示

分析結果に対し、チャット機能でインタラクティブに深掘りが可能

## 5. 業種別分析機能の環境構築および検証

### 5.6. 検証方法 検証手順

- 検証担当者がユースケース I・II に該当する機能を利用し、評価結果をアンケートに記載する。

#### ① 検証の実施

- 検証担当者が、ユースケース I・II に該当する機能を、アンケートに回答することを目的に利用する。

#### ② 検証アンケートの実施

- 総合評価で、検証全体を通しての検証全体アンケートと、ユースケース I・II それぞれに対するユースケース別アンケートを行う。
- 個別評価で、検証を実施した設定条件ごとにアンケートを行う。

№	アンケート項目名	観点	回答形式	必須/任意	説明概要	評価項目			
<b>検証全体アンケート</b>									
1	検証全体アンケート全体	継続利用の意思	選択式	必須	今回検証した際のAI連携機能（業種別分析機能）を引き続き利用したいと思うか。				
2	導入障壁	認知度	選択式	必須	AI連携機能（業種別分析機能）を知っているか。				
3		導入意欲	選択式	必須	AI連携機能（業種別分析機能）を導入するにあたって、業務統合や情報連携、システム連携などはどの程度まで実現する必要があるか。				
4		導入障壁の克服	選択式	任意	AI連携機能（業種別分析機能）の導入にあたって、どのような課題があるか。				
5		導入障壁の克服	選択式	必須	AI連携機能（業種別分析機能）の導入にあたって、どのような課題があるか。				
<b>業種別分析機能アンケート</b>									
6	業務効率化・品質向上	AI活用への期待	選択式	必須	抽出結果（業種別分析機能）の活用により、どのような効果が期待できるか。				
7		導入障壁	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
8		導入障壁の克服	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
9		導入障壁の克服	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
10		導入障壁の克服	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
11		導入障壁の克服	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
12	UI-UX	操作性	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
13		操作性	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
14		操作性	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
15		操作性	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
16		操作性	選択式	必須	抽出結果（業種別分析機能）の活用にあたって、どのような課題があるか。				
<b>評価記録</b>									
	#	機能設定	処理条件設定	事例区分設定	分析軸設定	AI回答の妥当性	体感応答速度	気づきの有無	問題点等備考
	1								
	2								
	3								
	4								
	5								
	6								

## 5. 業種別分析機能の環境構築および検証

### 5.6. 検証方法 アンケート詳細

- 検証担当者を対象に、以下の項目のアンケートを実施した。

アンケート種別		アンケート観点	アンケート観点概要
総合評価アンケート	検証全体アンケート	継続利用の意思	・ 業種別分析機能を引き続き利用したいか。
		導入障壁	・ 業種別分析機能の導入に障壁があるか。
		他業務への展開	・ 業種別分析機能をヒアリング記録以外の業務に利用できるか。
	ユースケース別アンケート	LLMの出力の質	・ LLMの出力内容が理解しやすく正しかったか。
		業務効率化・品質向上	・ 業種別分析機能により、業務の効率化や品質向上が実現したか。
		UI・UX	・ 業種別分析機能が操作しやすかったか。
		機能価値向上	・ 業種別分析機能の利用価値を向上する改善点があったか。
個別評価アンケート	レコード別アンケート	LLM回答の妥当性	・ 当該条件でLLMの出力内容が妥当であったか。
		体感応答速度	・ 当該条件でLLMの出力時間が許容できる範囲内であったか。
		気づきの有無	・ 当該条件で業種別分析機能を利用した中で改善点などの気づきがあったか。
		問題点等備考	・ 当該条件で業種別分析機能を利用した中で問題点があったか。

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 総合評価アンケート結果 検証全体アンケート

- 6割の検証者が業種別機能の継続利用を望んでいる。
- 利便性の低さの観点での課題がある。

アンケート概要	選択式アンケート結果	記述式アンケート結果										
<b>【継続利用の意思】</b> 業種別分析機能を引き続き利用したいと思うか。	<table border="1"><thead><tr><th>回答</th><th>件数</th></tr></thead><tbody><tr><td>1: とてもそう思う</td><td>0</td></tr><tr><td>2: そう思う</td><td>3</td></tr><tr><td>3: そう思わない</td><td>1</td></tr><tr><td>4: 全くそう思わない</td><td>1</td></tr></tbody></table>	回答	件数	1: とてもそう思う	0	2: そう思う	3	3: そう思わない	1	4: 全くそう思わない	1	<ul style="list-style-type: none"><li>■ 評価された点<ul style="list-style-type: none"><li>• 使い慣れれば便利である。</li><li>• 抽出や要約が優れている。</li></ul></li><li>■ 改善が期待される点<ul style="list-style-type: none"><li>• 抽出後にヒアリング結果を整理をするのが難しい。</li><li>• さらに絞り込んで抽出を行いたい。</li><li>• 出力結果が当然の内容であり、あえて本システムを使用するメリットがない。</li><li>• Power BIでの事例抽出と手間があまり変わらない。</li></ul></li></ul>
回答	件数											
1: とてもそう思う	0											
2: そう思う	3											
3: そう思わない	1											
4: 全くそう思わない	1											
<b>【導入障壁】</b> 業種別分析機能の導入への障壁があるか。	(選択式アンケートなし)	(改善点がある場合のみ回答) <ul style="list-style-type: none"><li>■ 改善が期待される点<ul style="list-style-type: none"><li>• 元のデータを確認する必要がある場合、二度手間に感じる。</li></ul></li></ul>										
<b>【他業務への展開】</b> 業種別分析機能をヒアリング記録以外の業務に利用できるか。	(選択式アンケートなし)	<ul style="list-style-type: none"><li>■ 他業務への展開案<ul style="list-style-type: none"><li>• kintoneの操作のヘルプ機能</li><li>• 関係法令に関するeラーニングシステム</li></ul></li></ul>										

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースⅠ）

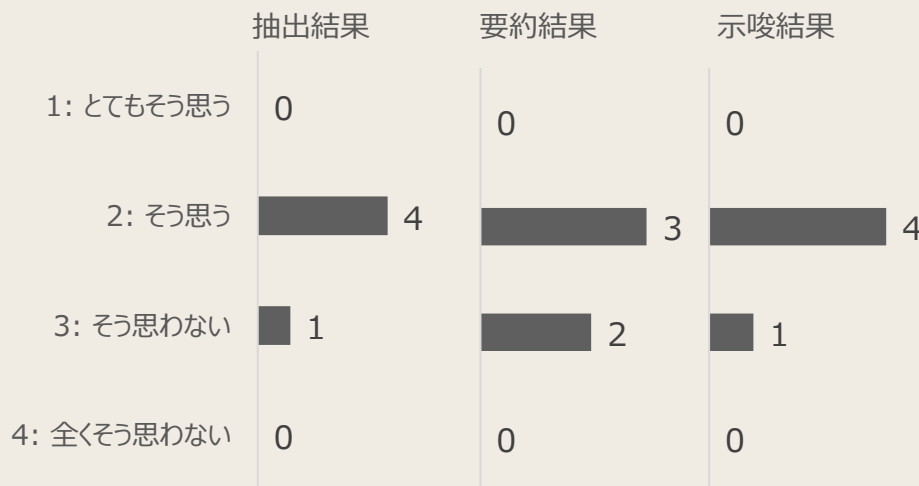
- 出力の質については、示唆の業務適応性の低さや業務固有の軸の考慮に関する課題がある。
- 業務効率化、品質向上については、利便性の低さに関する課題がある。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。

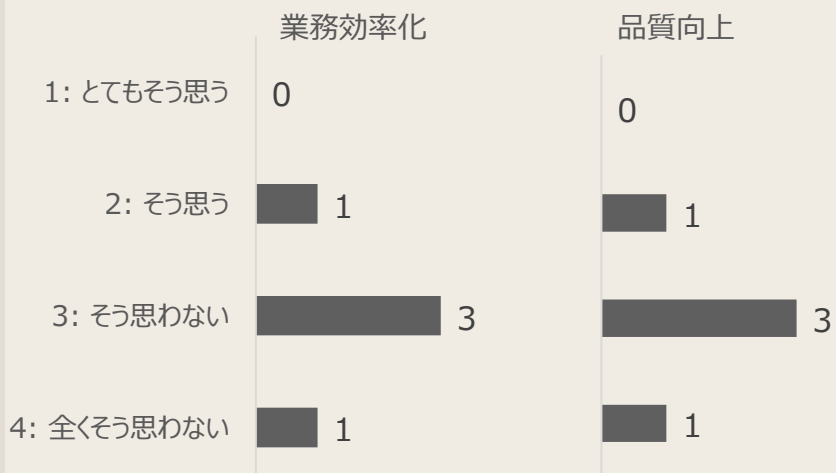


(改善点がある場合のみ回答)

■ 改善が期待される点

- 表示結果が事例の記述の一部であり、具体的な取引内容や事業内容等の情報が必要な場合がある。
- 示唆結果の用途がわからない。
- 一般的にはあっているが、業務特有の軸で見ると誤りである抽出結果がある。

**【業務効率化・品質向上】**  
業種別分析機能の導入への障壁があるか。



■ 評価された点

- 曖昧な抽出条件でも一定の抽出を行ってくれる。

■ 改善が期待される点

- もう少し手間が少なくないと使うメリットを感じない。
- 抽出の条件範囲が決められているため、より細かい抽出をしたい場合には使いづらい。
- エラーが出るため多くの件数を一気に抽出できない。

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースⅠ）

- UI・UXについては、条件設定や元データ確認、出力結果活用における利便性の低さに関する課題がある。

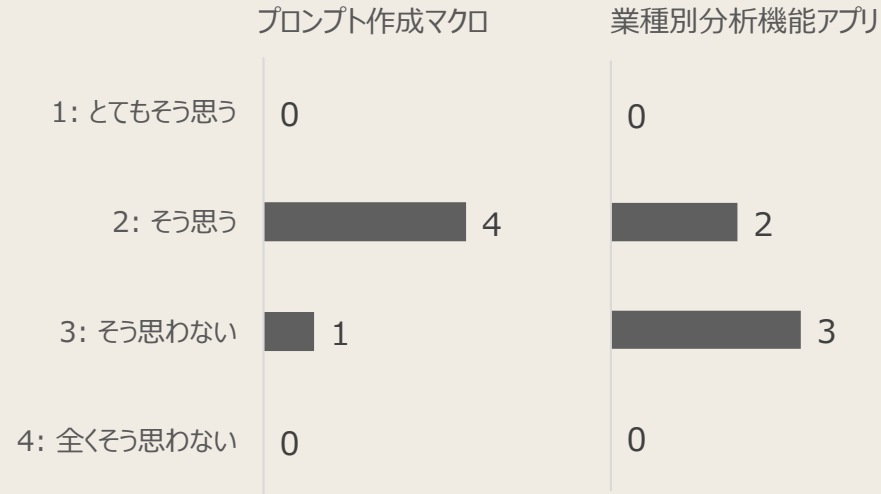
#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

##### 【UI・UX】

業種別分析機能が操作しやすいと思うか。



(改善点がある場合のみ回答)

##### ■ 改善が期待される点

- 複数の事例区分を選択できない。
- 登録ユニークキーがどのアプリのユニークキーかわからない。
- 出力結果からリンク等でkintoneに飛べるようにしてほしい。
- csvやエクセルに抽出結果をアウトプットできるようにしてほしい。

##### 【機能価値向上】

業種別分析機能に利用価値補  
向上する改善点があるか。

(選択式アンケートなし)

(改善点がある場合のみ回答)

##### ■ 改善が期待される点

- 企業名から事例を抽出する機能があるとよい。
- 抽出結果として、事例の法令等根拠を示すことはできないか。
- 条件の項目を増やして、より細かい条件をつけられるようにする方がよい。
- 条件の保存ができるようになるとうよい。

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースⅠ）

- 業務固有の軸を考慮すべき分析軸は複数存在すると考えられる。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

【LLM回答の妥当性】  
LLMの出力内容が妥当であるか。



（改善点がある場合のみ回答）

- 改善が期待される点
- 事例検索で別の分類軸の事例が抽出された。
- 細分課題を横断して該当する事例が抽出されている。
- 誤った事例が含まれる場合がある。

【体感応答速度】  
LLMの出力にかかる時間が許容できる範囲内か。



（改善点がある場合のみ回答）

- 改善が期待される点
- （具体的なコメントなし）

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースⅡ）

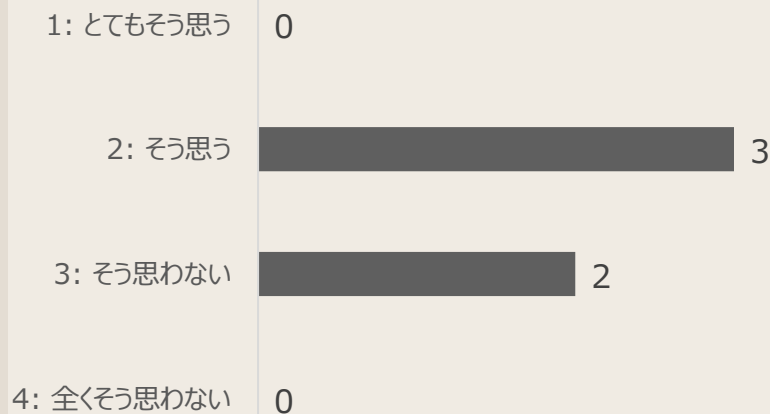
- 出力の質については、業務固有の軸の考慮に関する課題がある。
- 業務効率化、品質向上については、利便性の低さに関する課題がある。

#### アンケート概要

#### 選択式アンケート結果

#### 記述式アンケート結果

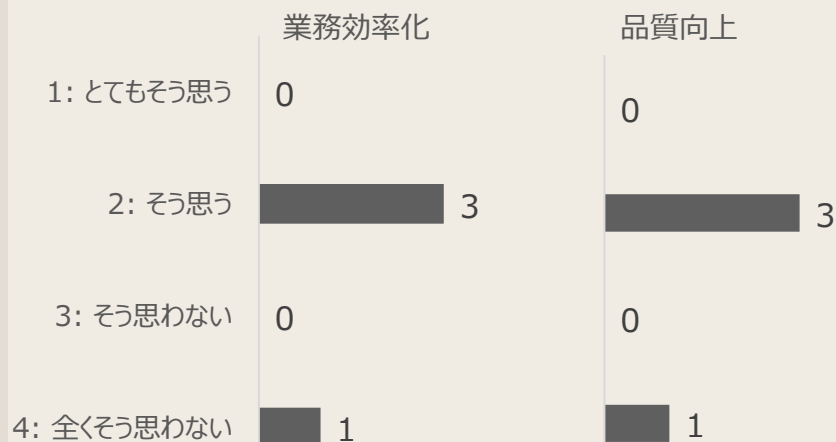
**【LLMの出力の質】**  
LLMの出力内容が理解しやすく正しいと思うか。



(改善点がある場合のみ回答)

- 改善が期待される点
  - ・ 一般的にはあっているが、業務固有の軸で見ると誤りである抽出結果がある。
  - ・ 誤りはないが、出力される情報が当然の内容であり、積極的に利用したいと思わない。

**【業務効率化・品質向上】**  
業種別分析機能の導入への障壁があるか。



- 評価された点
  - ・ 任意のキーワードでヒアリング結果の抽出ができるのは良いと思う。
  - ・ 曖昧な抽出条件でも一定の抽出を行ってくれる。
- 改善が期待される点
  - ・ power BIでの事例抽出と手間があまり変わらない。

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 総合評価アンケート結果 ユースケース別アンケート（ユースケースⅡ）

- UI・UXについては、条件設定や元データ確認、出力結果活用における利便性の低さに関する課題がある。

アンケート概要	選択式アンケート結果	記述式アンケート結果															
<b>【UI・UX】</b> 業種別分析機能が操作しやすいと思うか。	<table border="1"><thead><tr><th></th><th>プロンプト作成マクロ</th><th>業種別分析機能アプリ</th></tr></thead><tbody><tr><td>1: とてもそう思う</td><td>0</td><td>0</td></tr><tr><td>2: そう思う</td><td>4</td><td>2</td></tr><tr><td>3: そう思わない</td><td>1</td><td>1</td></tr><tr><td>4: 全くそう思わない</td><td>0</td><td>1</td></tr></tbody></table>		プロンプト作成マクロ	業種別分析機能アプリ	1: とてもそう思う	0	0	2: そう思う	4	2	3: そう思わない	1	1	4: 全くそう思わない	0	1	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• 分析軸が一つしか選定できないので、何度もやり直す必要がある。</li><li>• 出力結果からリンク等でkintone1に飛べるようにしてほしい。</li><li>• csvやエクセルに抽出結果をアウトプットできるようにしてほしい。</li><li>• 複数の事例区分を選択できない。</li></ul>
	プロンプト作成マクロ	業種別分析機能アプリ															
1: とてもそう思う	0	0															
2: そう思う	4	2															
3: そう思わない	1	1															
4: 全くそう思わない	0	1															
<b>【機能価値向上】</b> 業種別分析機能に利用価値補向上する改善点があるか。	(選択式アンケートなし)	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• 条件ごとに「OR」か「AND」か選べるようにした方がよい。</li><li>• 条件の保存ができるようになるとよい。</li></ul>															

## 5. 業種別分析機能の環境構築および検証

### 5.7. 検証結果 個別評価アンケート結果 ユースケース別アンケート（ユースケースⅡ）

- 業務固有の軸を考慮すべき分析軸は複数存在すると考えられる。

アンケート概要	選択式アンケート結果	記述式アンケート結果								
<b>【LLM回答の妥当性】</b> LLMの出力内容が妥当であるか。	<table border="1"><thead><tr><th>評価</th><th>回数</th></tr></thead><tbody><tr><td>1: 良好</td><td>0</td></tr><tr><td>2: やや不良</td><td>1</td></tr><tr><td>3: 不良（誤答）</td><td>1</td></tr></tbody></table>	評価	回数	1: 良好	0	2: やや不良	1	3: 不良（誤答）	1	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• 想定と異なる分析軸の内容が抽出された。</li><li>• 誤った事例が含まれる場合がある。</li></ul>
評価	回数									
1: 良好	0									
2: やや不良	1									
3: 不良（誤答）	1									
<b>【体感応答速度】</b> LLMの出力にかかる時間が許容できる範囲内か。	<table border="1"><thead><tr><th>評価</th><th>回数</th></tr></thead><tbody><tr><td>1: 許容範囲内</td><td>2</td></tr><tr><td>2: 運用上不適</td><td>0</td></tr></tbody></table>	評価	回数	1: 許容範囲内	2	2: 運用上不適	0	<p>(改善点がある場合のみ回答)</p> <ul style="list-style-type: none"><li>■ 改善が期待される点</li><li>• (具体的なコメントなし)</li></ul>		
評価	回数									
1: 許容範囲内	2									
2: 運用上不適	0									

## 5. 業種別分析機能の環境構築および検証

### 5.8. 今後対応すべき課題

- アンケート結果から品質・UXに関する6つの課題を洗い出した。
- 各課題に対して対応策を検討する必要がある。

課題名	課題名	課題概要	対応アンケートコメント例	ユースケース
意思決定支援としての品質	示唆出しの精度向上	<ul style="list-style-type: none"> <li>LLMによる示唆出しの出力結果が、業務活用レベルに至っていない</li> </ul>	<ul style="list-style-type: none"> <li>示唆結果の用途がわからない。</li> </ul>	I
	業務固有の分析軸考慮	<ul style="list-style-type: none"> <li>業務固有の軸を踏まえた細分課題レベルでの抽出が実施できておらず、業務活用に至っていない（他業務の固有軸と混在するケースがある）</li> </ul>	<ul style="list-style-type: none"> <li>一般的には合っているが、業務固有の軸で見ると誤りである抽出結果がある。</li> <li>業務固有の細分課題を横断して該当する事例が抽出されている。</li> <li>事例検索で別の分類軸の事例が抽出された。</li> </ul>	I / II
システムにおけるUX	出力結果の円滑な業務利用	<ul style="list-style-type: none"> <li>出力結果の根拠の妥当性を効率的に確認できず、またデータのエクスポートができないため、円滑に業務利用ができない</li> </ul>	<ul style="list-style-type: none"> <li>元のデータを確認する必要がある場合、二度手間を感じる。</li> <li>出力結果からリンク等でkintoneに飛べるようにしてほしい。</li> <li>表示結果が事例の記述の一部であり、具体的な取引内容や事業内容等の情報が必要な場合がある。</li> <li>csvやエクセルに抽出結果をアウトプットできるようにしてほしい。</li> <li>抽出後にヒアリング結果を整理するのが難しい。</li> </ul>	I / II
	条件設定機能の柔軟性	<ul style="list-style-type: none"> <li>データ抽出の条件設定は、多岐にわたる業種や分析目的に応じた利活用において柔軟性がない</li> </ul>	<ul style="list-style-type: none"> <li>さらに絞り込んで抽出を行いたい。</li> <li>抽出の条件範囲が決められているため、より細かい抽出をしたい場合には使いづらい。</li> <li>複数の事例区分を選択できない。</li> <li>企業名から事例を抽出する機能があるとよい。</li> <li>条件の保存ができるようになるとうい。</li> </ul>	I / II
	アプリ制約・品質	<ul style="list-style-type: none"> <li>LLMやシステムの技術的な制約により抽出数の上限があり、分析業務の範囲が制限されている</li> </ul>	<ul style="list-style-type: none"> <li>エラーが出るため多くの件数を一気に抽出できない。</li> </ul>	I / II

## 6. 今後の取り組み

6.1. 今後想定される課題と取り組み

6.2. 取引調査員におけるヒアリング業務のロードマップ

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 入力補助機能

- 入力補助機能の検証における、類型化した課題を品質、UX、運用の観点で整理した。各類型について対策案の検討方針を示す。

課題の類型と概要			対策案の検討方針
意思決定支援としての品質	入力元情報の品質向上	LLMの入力元となるデータの質または量が十分でなく、分類などの出力結果の精度が上がらない	<ul style="list-style-type: none"> <li>分類に用いる入力元情報の精度を上げる処理を施し、意思決定支援に資する精度向上を目指す</li> </ul>
	優先順位付け	LLMが提示した候補が複数ある場合、職員はその全てを吟味する必要があり、確認工数が大きい	<ul style="list-style-type: none"> <li>最有力候補の提示や優先順位付け手法を検討し、職員の判断負荷低減を図る</li> </ul>
	文章標準化の品質向上	LLMで生成した標準化文章が、業務に適用可能なレベルに至っていない	<ul style="list-style-type: none"> <li>出力前の校閲結果をLLMで再評価を行うことで、文章標準化の精度向上を図る</li> </ul>
	専門知識の考慮	専門用語を含む文章の校閲において、専門的な知識が必要となる処理を適切に行えていない	<ul style="list-style-type: none"> <li>専門用語集などを整備・活用し、専門用語を適切にフィルタリング・修正する手法を検討</li> </ul>
	法令基準の反映	法令で定められた基準に基づくカテゴリ判定が十分に行えていない	<ul style="list-style-type: none"> <li>法令で定められた基準に沿ったロジック・プロセスを整理・検討する</li> </ul>
システムにおけるUX	応答時間の短縮	LLMを含むシステムの処理時間が長く、業務効率化および円滑な利用を妨げている	<ul style="list-style-type: none"> <li>処理状況の透明性を向上させ、職員の不安と体感待ち時間を軽減する</li> </ul>
	システムのエラー耐性向上	機能が正常に動作せず、想定した利用シナリオで機能を利用できないケースがある	<ul style="list-style-type: none"> <li>各システムエラーを調査し、継続的な改善をすることで、職員のシステムに対する信頼を獲得する</li> </ul>
運用・ガバナンス（人材育成）	専門性・判断力のスキルアップ	LLMへの過度な依存は、職員の専門性低下やヒアリング記録の品質低下を引き起こすリスクがある	<ul style="list-style-type: none"> <li>単に利便性を向上させるだけでなく、共有プラットフォーム構築や継続的な教育の検討が必要</li> </ul>

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 業種別分析機能

- 業種別分析機能の検証における、類型化した課題を品質、UXの観点で整理した。各類型について対策案の検討方針を示す。

課題の類型と概要		対策案の検討方針
意思決定支援としての品質	示唆出しの精度向上	LLMによる示唆出しの出力結果が、業務活用レベルに至っていない
	業務固有の分析軸考慮	業務固有の軸を踏まえた細分課題レベルでの抽出が実施できておらず、業務活用に至っていない（他業務の固有軸と混在するケースがある）
システムにおけるUX	出力結果の円滑な業務利用	出力結果の根拠の妥当性を効率的に確認できず、またデータのエクスポートができないため、円滑に業務利用ができない
	条件設定機能の柔軟性	データ抽出の条件設定は、多岐にわたる業種や分析目的に応じた利活用において柔軟性がない
	アプリ制約・品質	LLMやシステムの技術的な制約により抽出数の上限があり、分析業務の範囲が制限されている



- より広範なユースケースにおける有用な示唆出しの要件を整理した上で、**プロンプト調整等**を検討・実施



- 業務固有の細分課題を**業務要件レベルで絞込みし**、業務固有の軸の**切り分けを厳格にするロジック**を検討



- LLM回答の**根拠の妥当性を迅速に確認可能**とした上で、資料作成や**後続業務への活用**を考慮



- 絞込み機能の**UI/UX向上**および**目的に沿った分析を両立**する対策案を検討



- LLMやシステムの技術的な制約事項を整理し、**業務要件との擦り合わせ**を精緻に行い、**LLMモデルの変更**なども視野に追加検討を実施

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 入力元情報の精緻化による出力の精度向上

入力補助機能

意思決定支援  
としての品質

入力元情報  
の品質

課題の類型と概要

LLMの入力元となるデータの質または量が十分でなく、分類  
などの出力結果の精度が上がらない



対策案の検討方針

- 分類に用いる入力元情報の精度を上げる処理を施し、意思決定支援に資する精度向上を目指す

#### 入力元情報の精緻化プロセス案

入力元情報



LLMが入力情報で  
分類可能か判断



結果を出力



※既存の入力情報および新規の入力情報

追加情報の入力

入力情報  
+ 追加情報

不足情報の追加入力をユーザに促す  
(例：〇〇観点について詳しく記載ください。)

(追加) 入力情報の精緻化プロセス

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 抽出結果の優先順位付け

入力補助機能

意思決定支援  
としての品質

優先順位付け

課題の類型と概要

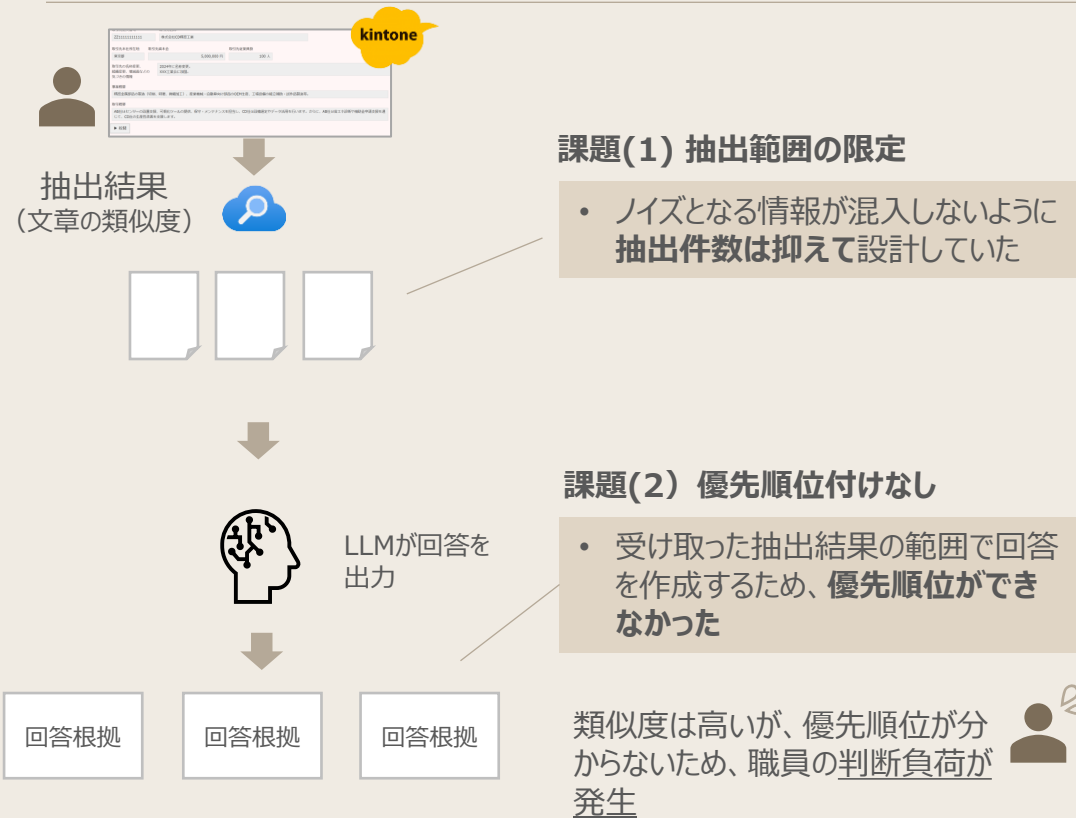
LLMが提示した候補が複数ある場合、職員はその全てを吟味する必要があり、確認工数が多い



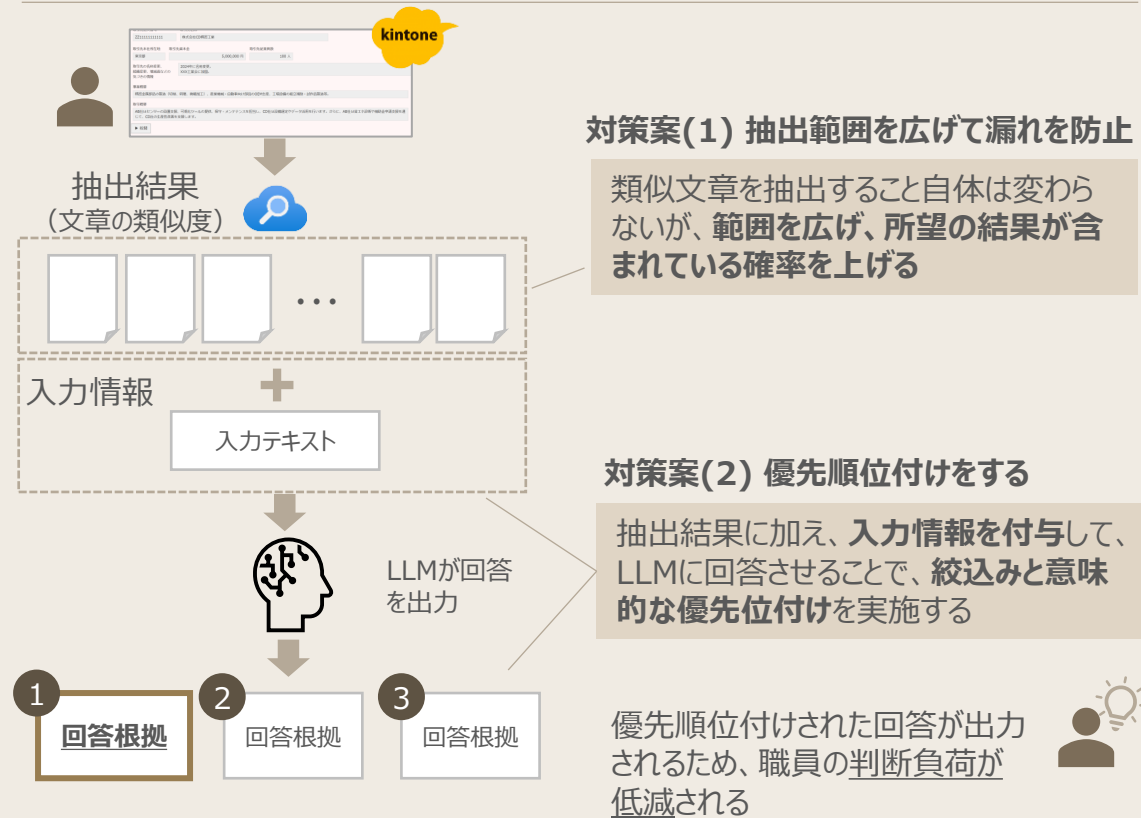
対策案の検討方針

- 最有力候補の提示や優先順位付け手法を検討し、職員の判断負荷低減を図る

現状の抽出の仕組み



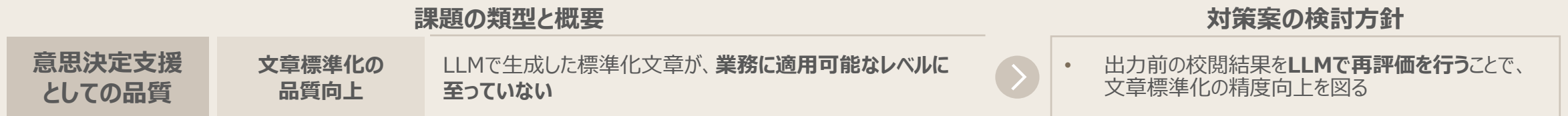
抽出結果の優先順位付け案



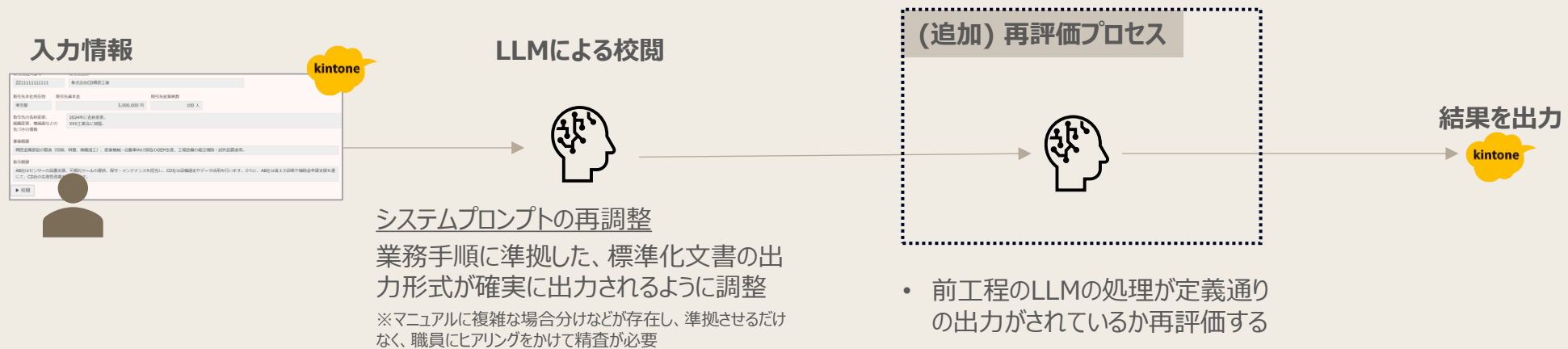
## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み LLMによる校閲結果の再評価

入力補助機能



### 再評価プロセスの追加案



## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 専門用語の誤修正防止

入力補助機能

#### 課題の種類と概要

意思決定支援  
としての品質

専門知識の考慮

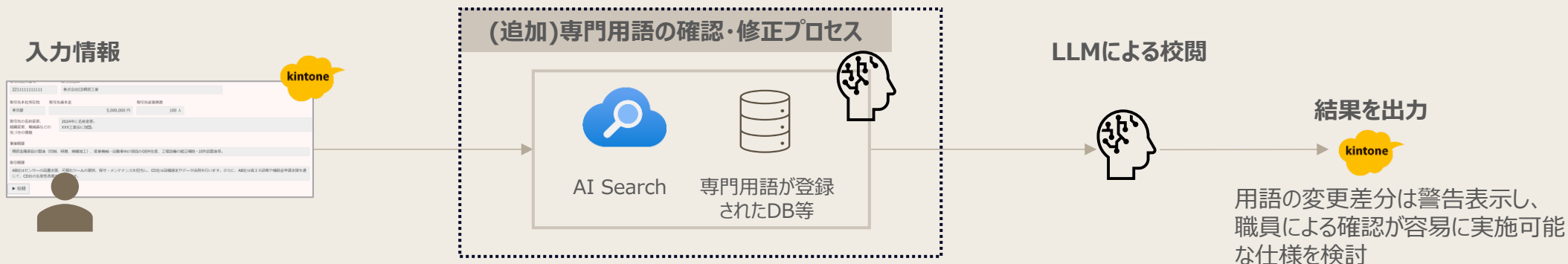
専門用語を含む文章の校閲において、**専門的な知識が必要となる処理を適切に行えていない**



#### 対策案の検討方針

- 専門用語集などを整備・活用し、**専門用語を適切にフィルタリング・修正する手法**を検討

#### 専門用語の確認・修正プロセス案



#### 専門用語の修正方針

- ✓ 専門用語集等のDBに登録されている用語と、**類似しているが異なる用語**をLLMによって検知した場合、DB登録情報で修正を行う
- ✓ 一方、**正しく使用されている専門用語**については誤って変更させない処理を施す

<前提>

- 専門用語集などの整備が必要
- 用語集の保守運用方法の検討も必要

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 業務固有の判断プロセス

入力補助機能

意思決定支援  
としての品質

法令基準の反映

課題の類型と概要

法令で定められた基準に基づくカテゴリ判定が十分に行えていない



対策案の検討方針

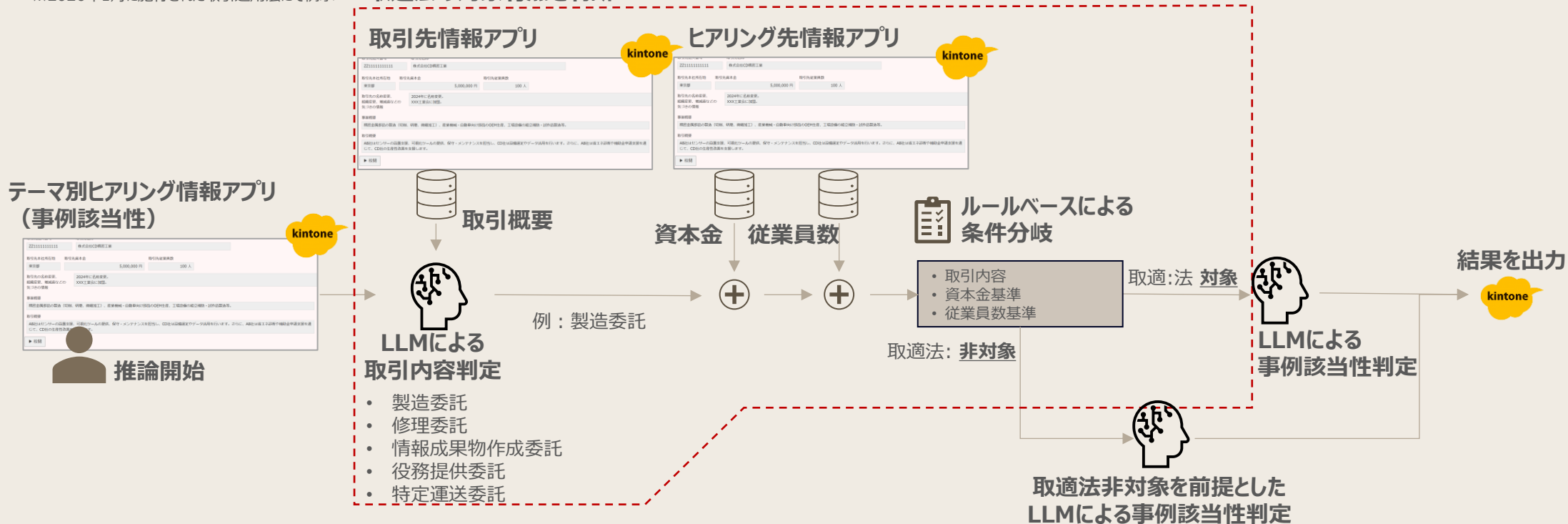
- 法令で定められた基準に沿ったロジック・プロセスを整理・検討する

### 企業情報に則った判断プロセス案

例：事例該当性

※2026年1月に施行された取引適用法にて例示

LLMで事例該当性を判断の前に、ルールベースで取引内容・資本金基準・従業員数基準を基に  
取適法の対象有無を判断



## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 処理工程の可視化による利用体験向上

入力補助機能

#### 課題の種類と概要

システムにおけるUX

応答時間の短縮

LLMを含むシステムの処理時間が長く、業務効率化および円滑な利用を妨げている



#### 対策案の検討方針

- 処理状況の透明性を向上させ、職員の不安と体感待ち時間を軽減する

#### kintone画面にステータス表示

処理工程を示す画面は、ユーザーが処理開始のボタンを押下した際に表示

#### kintone画面



#### 処理工程の可視化案

現時点の処理工程を表示することで職員の**体感待ち時間を軽減**

(ステータス表示例)

検索開始

- 検索の開始を伝える

検索中  
(進捗: 3/10)

- 検索中であることを伝える  
※可能であれば、件数の表示も検討することで、体感待ち時間を更に軽減

LLMによる  
回答を生成中..

- LLMによる回答中で、最終の処理工程であることを伝える

職員に回答を表示

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 根拠確認の負荷低減および分析結果のエクスポート

業種別分析機能

#### 課題の類型と概要

システムにおけるUX

出力結果の円滑な業務利用

出力結果の根拠の妥当性を効率的に確認できず、またデータのエクスポートができないため、円滑に業務利用ができない



#### 対策案の検討方針

- LLM回答の妥当性を迅速に確認可能とした上で、資料作成や後続業務への活用を考慮

#### LLM出力結果UI案

分析軸：支払条件に関する事例(好事例) LLMの出力画面

抽出結果

事例1

・ 締日である月末日に検収を受け、同日に請求額を確認し、翌月末日に100%現金にて支払を受けている。

参照情報： [取適法 XX条](#)

項目	内容
調査日	2025-XX-XX
担当局	XXXX経済産業局
登録ユニークキー	<a href="#">1234556789XX</a>
ヒアリング先	WXYZ 株式会社 (法人番号: 5678123A49876XX / 業種: ZZZ業)
取引先	ABCD 株式会社 (法人番号: 0987654321XX / 業種: AAA業)

事例2

メッセージを入力して [Enter] で送信、[Shift] + [Enter] で改行します。

[分析結果を出力](#)

#### 出力の根拠提示

LLMの出力結果に対する根拠を提示し、出力の妥当性を高める。また、職員が根拠内容を迅速に確認することを容易にすることを旨とする。

(例)

- 参照した法令の該当箇所を表示
- 登録ユニークキー\*のリンクより対象のkintoneレコード情報を取得・表示

#### 分析結果のファイル出力

- 職員が資料作成を円滑に行うため、出力情報をCSVやExcel形式等でのファイル出力を検討する。

(仮説) LLMやエージェント等活用による資料作成までの自動化

将来的な取組の方向性の一案として、所望の資料形式への取りまとめや体裁を含めた整形・出力までをLLMやエージェント等で実現可能かについて検討する余地があると考える。その結果次第では、職員の分析業務の高度化や業務効率の向上に資する可能性がある。

\*登録ユニークキー：各kintoneアプリに割り振られたレコードID

## 6. 今後の取り組み

### 6.1. 今後想定される課題と取り組み 絞込みの柔軟性向上

業種別分析機能

#### 課題の種類と概要

システムにおけるUX

条件設定機能の柔軟性

データ抽出の条件設定は、多岐にわたる業種や分析目的に応じた利活用において柔軟性がない



#### 対策案の検討方針

- 絞込み機能のUI/UX向上および目的に沿った分析を両立する対策案の検討

#### 業種別分析機能における検証結果

アンケート調査の結果より、以下の業務上のニーズを確認

- 詳細な抽出条件設定
  - 抽出の条件範囲が決められているため、より細かい抽出をしたい場合には使いづらい（本事業では、LLMのコンテキストウィンドウ等を考慮して期間等の範囲を限定した）
  - 複数の事例区分を選択できない
  - “AND”や“OR”か選べるようにした方がよい など
- キーワード検索
  - キーワードを自由入力して抽出できるとよい
- 条件設定の繰り返し利用
  - 条件の保存ができるとよい

より細かい条件設定による、詳細な抽出が業務要件として存在

#### 絞込みの柔軟性向上に係る対策案

- ✓ 期間・事例区分など詳細絞込み条件の追加
- ✓ AND/ORの検索条件の追加
- ✓ キーワードの自由入力検索による絞込み
- ✓ 条件の保存・呼出し

絞込み機能がリッチになるにつれ、条件設定が煩雑になり  
職員の分析業務におけるUX低下が予想される

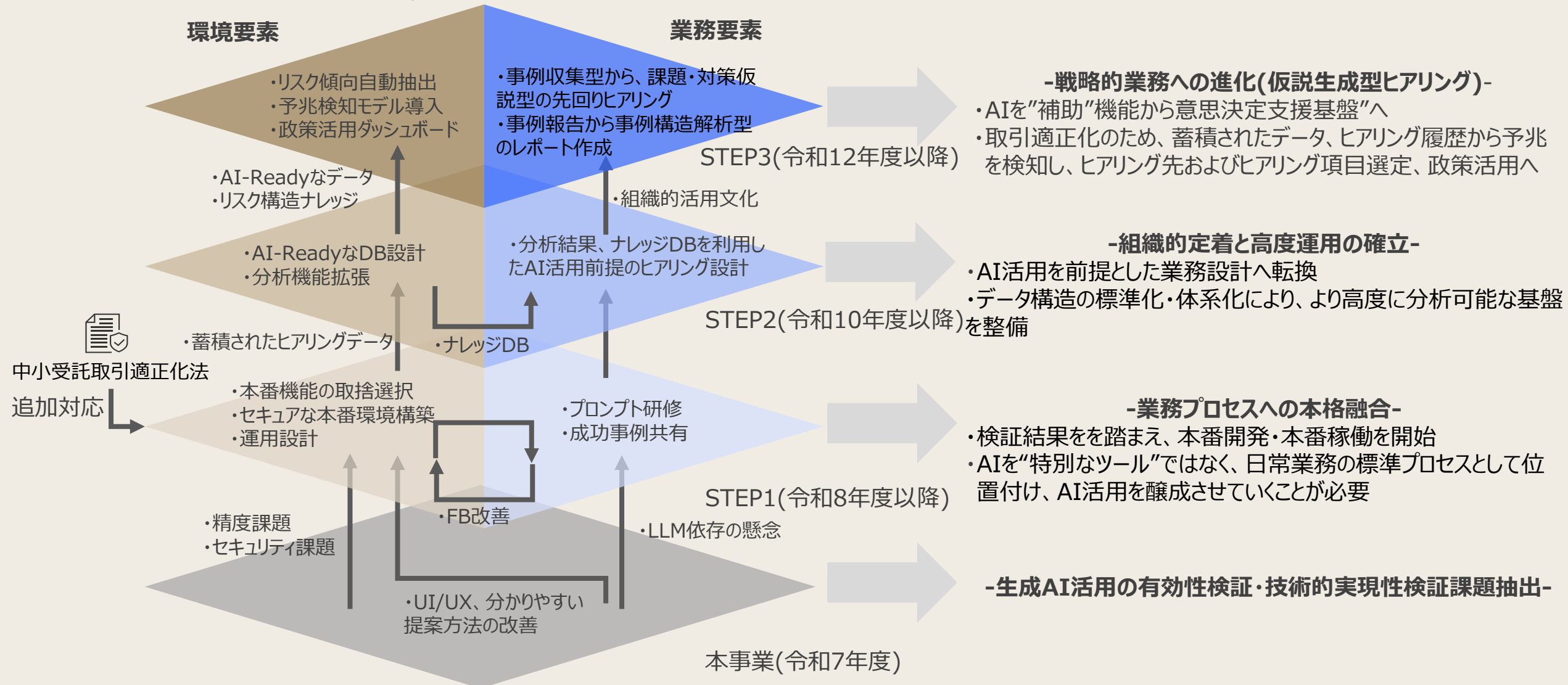
（仮説）LLMによる最適な抽出条件決定・分析をする業務高度化

- 分析の内容ごとに職員によるテンプレート等、よく利用される絞込み条件や観点を蓄積していき、抽出・分析の効率化に活かす仕組み
- 将来的な取組として、職員が分析要件をLLMと会話していく過程で、最適な条件を設定し、必要な情報をLLMが適時抽出・分析

## 6. 今後の取り組み

### 6.2. 取引調査員におけるAIを活用したヒアリング業務に向けたロードマップ（提案）

- 標準化されたヒアリングデータを蓄積、そのデータからリスク仮説を予知することで、事例収集型のヒアリング・分析・提言から、仮説確認型のヒアリング・提言といった戦略的業務への進化が期待される。



# Build Beyond As One.®



アビーム、ABeamおよびそのロゴは、アビームコンサルティング株式会社の日本その他の国における登録商標です。  
本文に記載されている会社名および製品名は各社の商号、商標又は登録商標です。