

トピックモデリングを用いたLLMによる要件の自動矛盾検知の手法

オープンソースソフトウェア工学研究室
S2310236 古川陽希

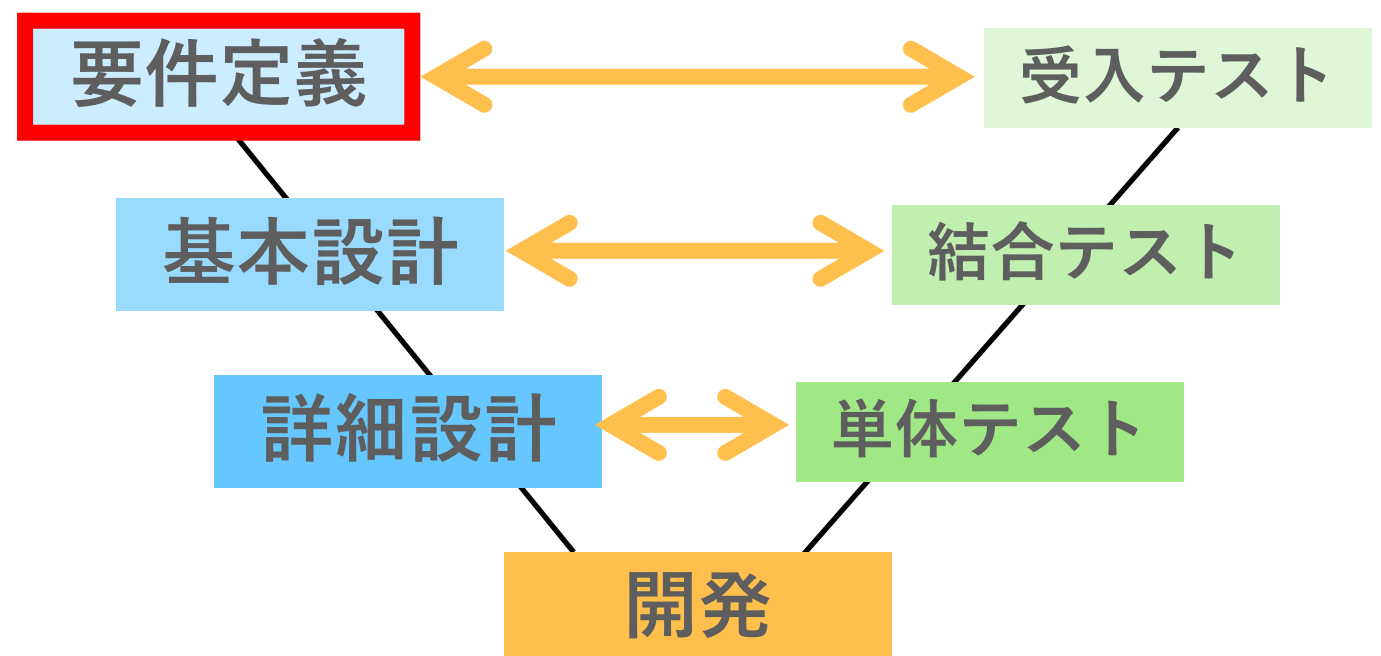
要件定義

システムが備えるべき機能や性能を明確に定義する工程



必要性

- 開発における最初のステップ
 - 欠陥発見⇒手戻り大
 - 開発者のゴールの明確化



要件定義の現状

要件同士の間隔が無い必要がある

①通信障害時や電波の届かない店舗であっても決済を行えるよう、**オフライン状態**での決済処理を許可し、完了できなければならない。

②残高の改ざんや二重支払いを確実に防止するため、すべての決済処理において必ず**オンライン通信**を行い、サーバー側の承認を得る必要がある。

⇒×

手動による意味論的な分析が必要

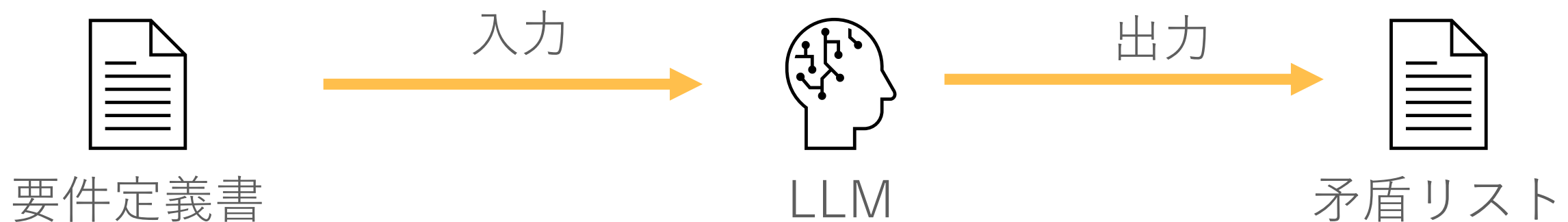
人間の専門家が行っていた

- ・膨大なコスト
- ・非常に難しいタスク

⇒ **自動矛盾検知**によって負担を軽減

自動矛盾検知と課題

要件の総当たり



大規模な要件定義書の処理

- ・ 比較回数が**多すぎる**

例) 200個の要件の計算量 = $\frac{200(200-1)}{2} = 19900$ 回

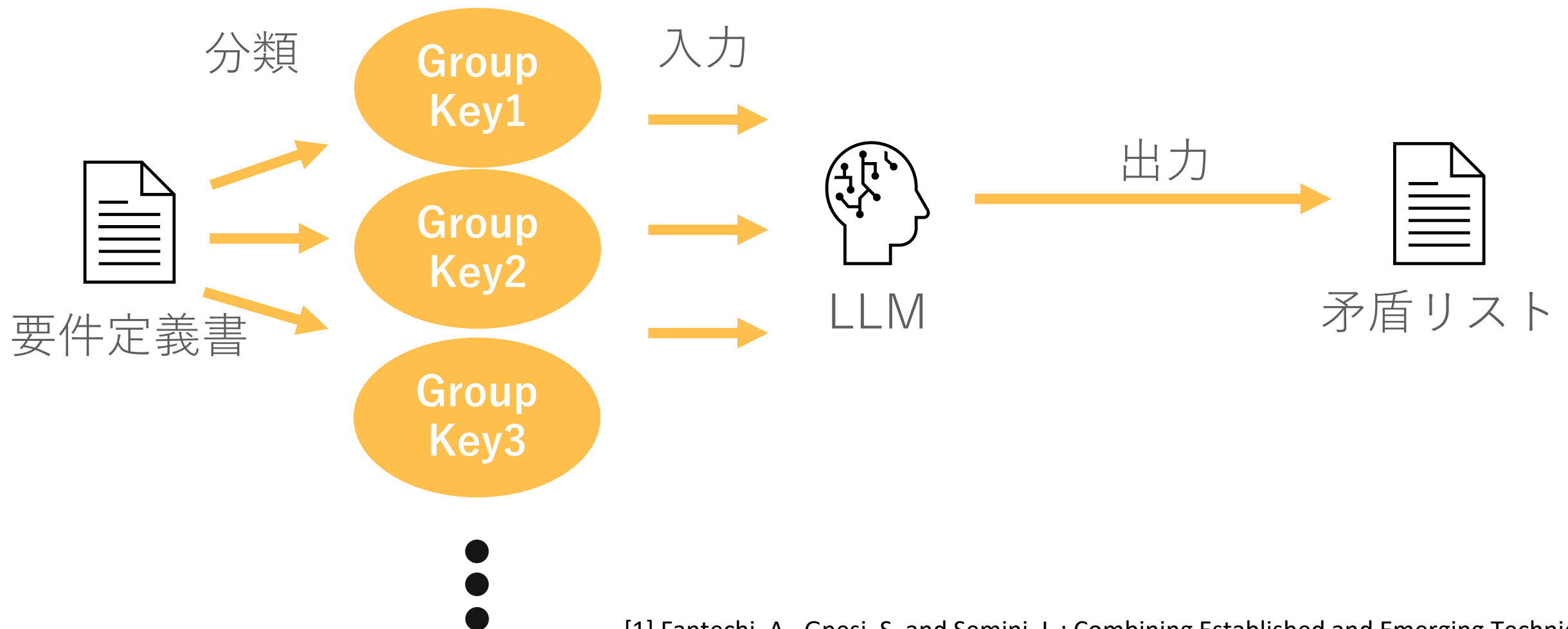
⇒ LLMの精度が低下

- ・ 出力をセーブしてしまう

先行研究

ルールベースによるクラスタリング^[1]

- ①キーワードをあらかじめ人が決定
- ②キーワードを含む要件をグループに分類する
- ③グループごとにLLMが矛盾検知



[1] Fantechi, A., Gnesi, S. and Semini, L.: Combining Established and Emerging Techniques to Detect Inconsistencies in Requirements, In Proceedings of RE '25, pp.519–526 (2025)

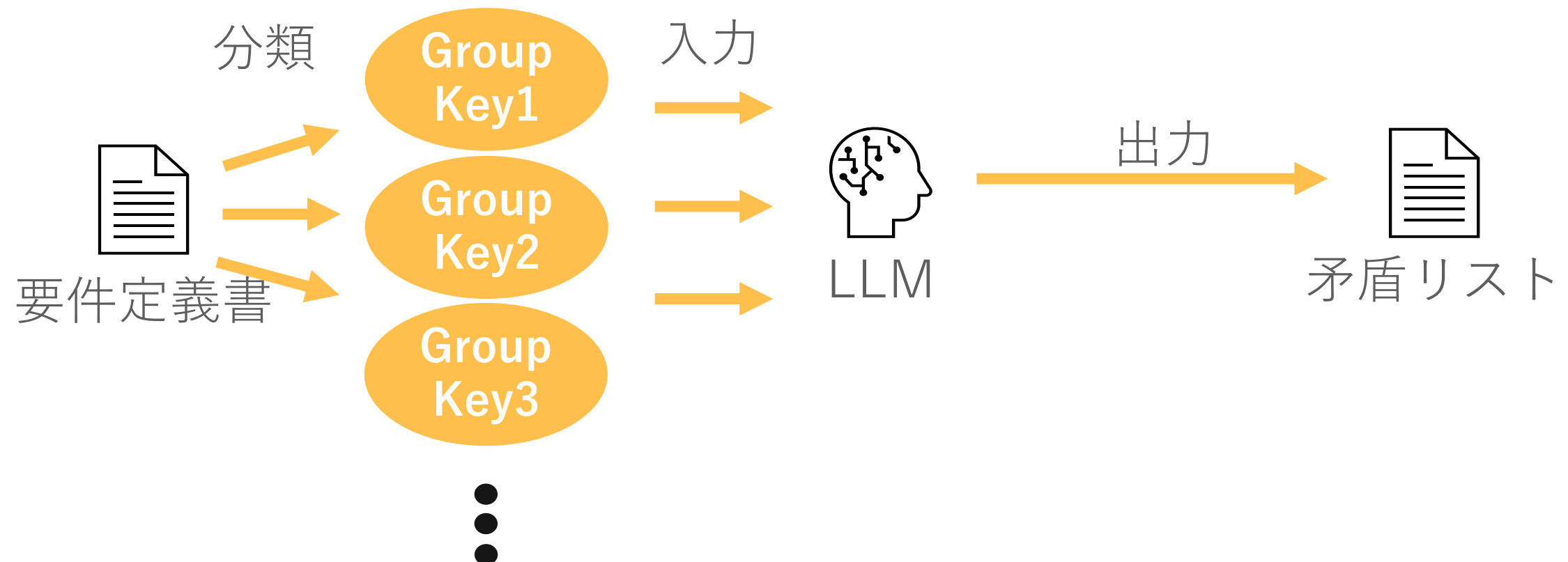
クラスタリングによる自動矛盾検知

仮定

- ・似ている内容の要件同士に矛盾

グループごとにLLMが矛盾を検知

- ・一回の比較回数が劇的に減少

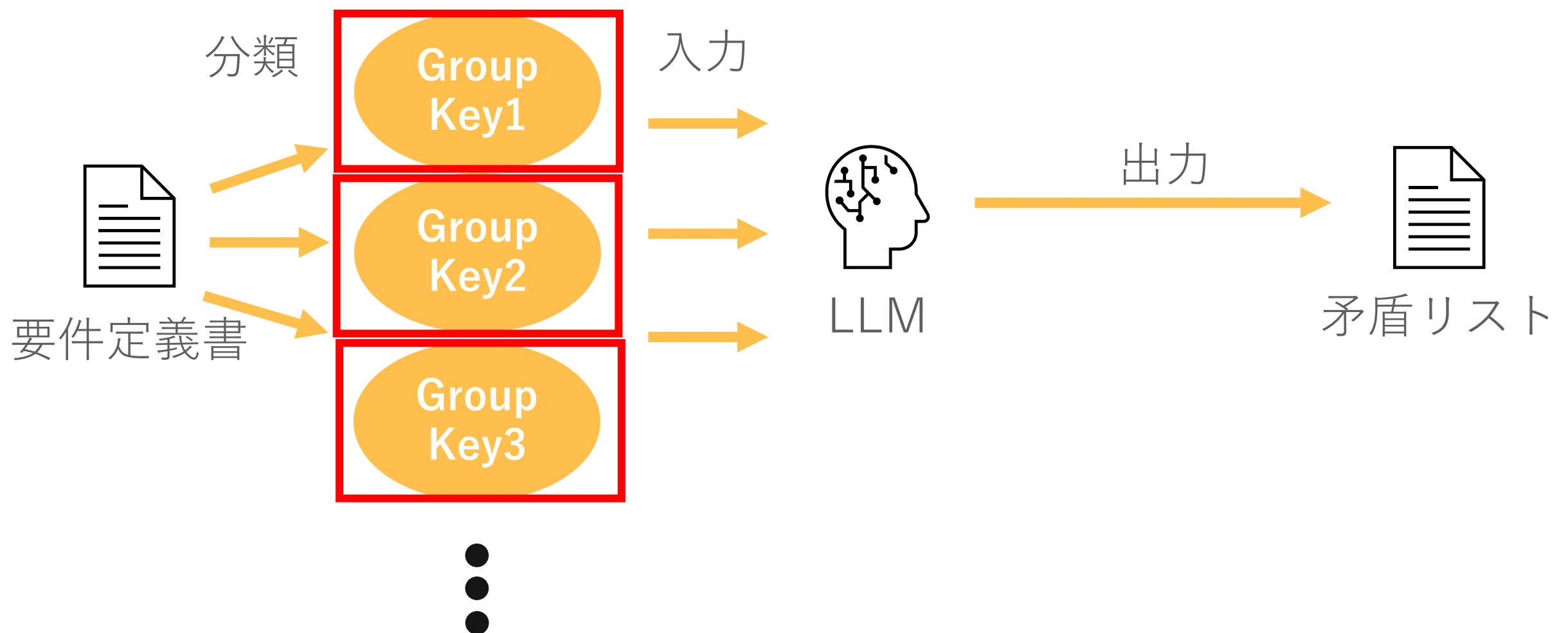


先行研究の課題

ルールベースによるクラスタリング

×それぞれの要件定義書で人間がキーワードを決める必要がある

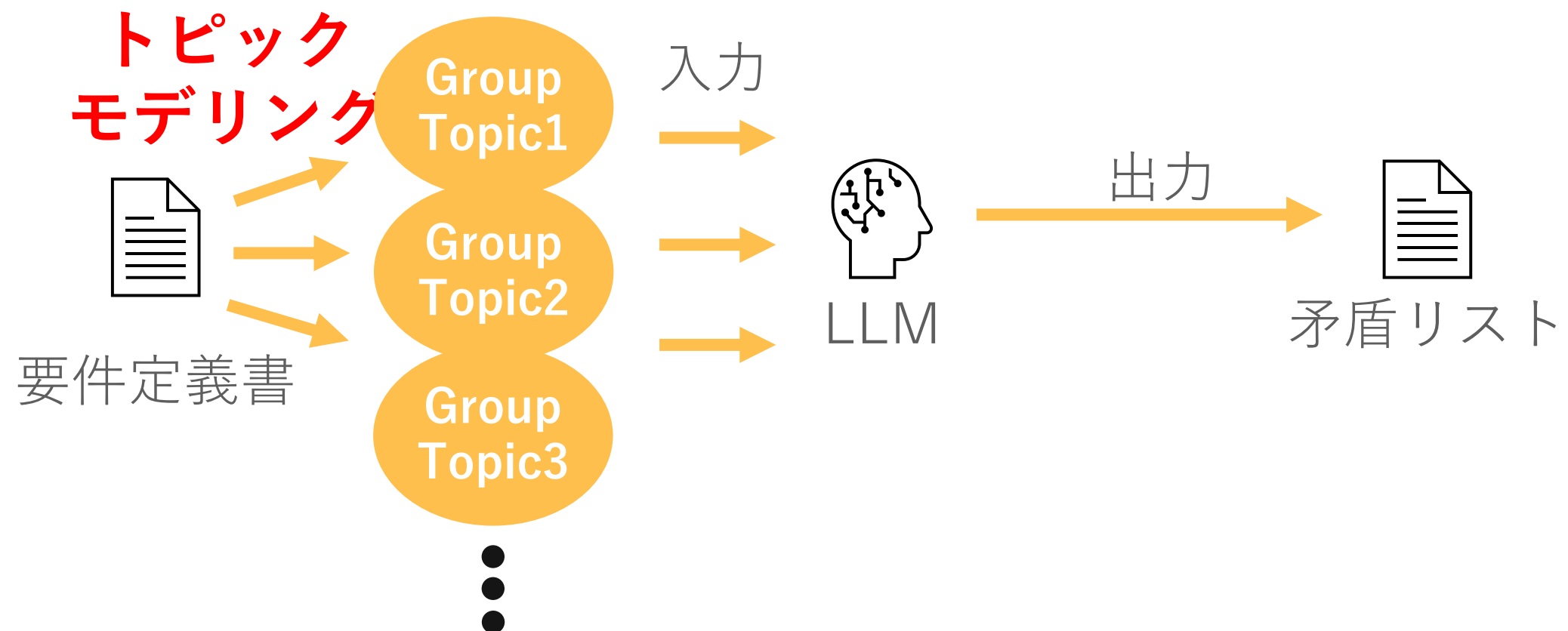
人間がキーワードを決定する



本研究について

要件に対して**トピックモデリング（BERTopic）**を行って分類

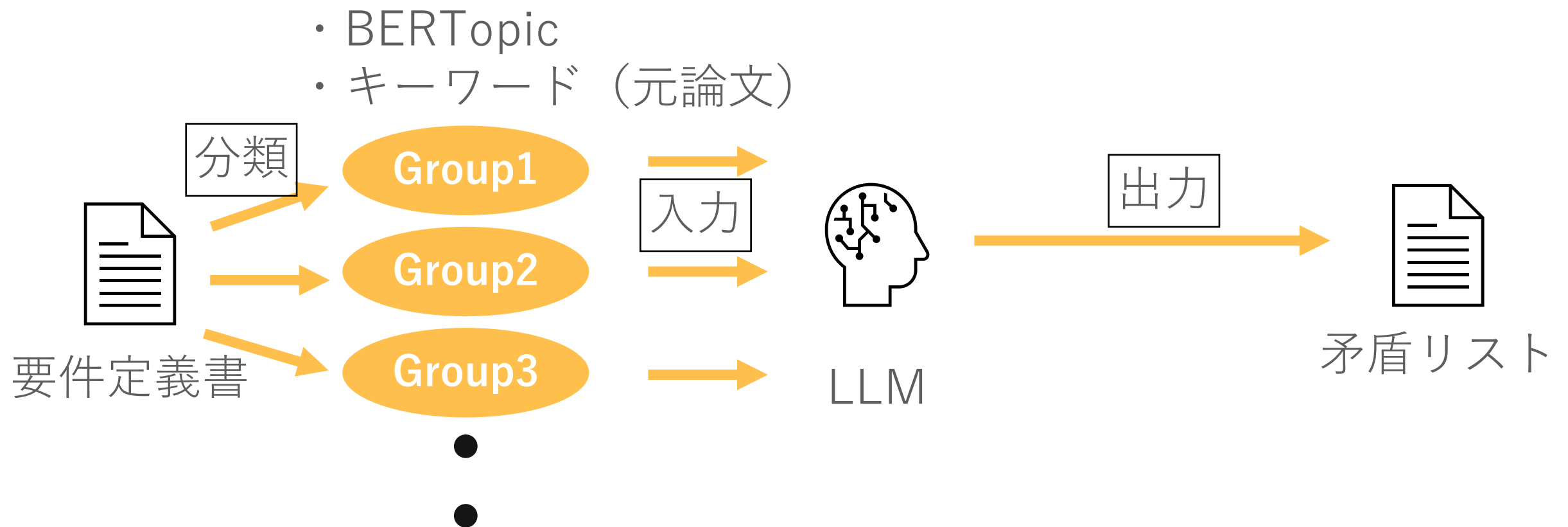
- ・ **教師無し学習**：潜在的なトピックを自動抽出
- ⇒ トピックを人間が考える必要がない



実験

実験手順

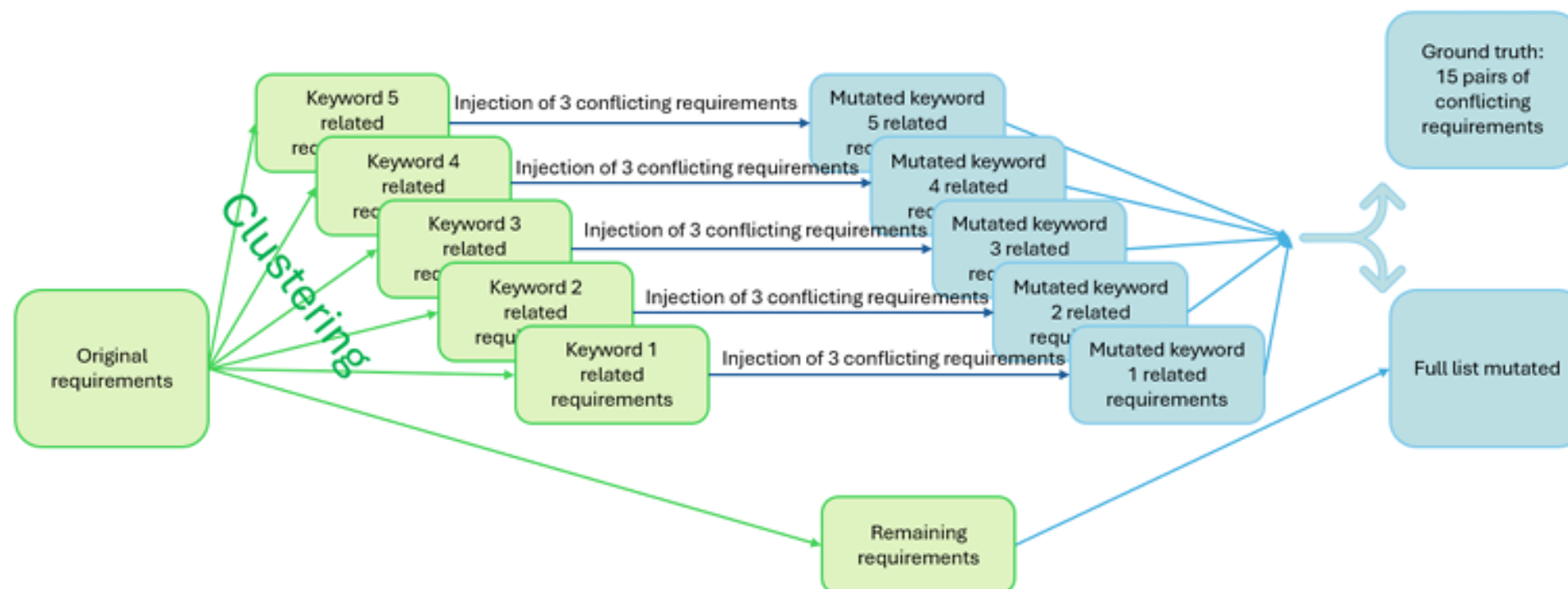
- ① 先行研究の論文で使用されたバグを仕込んだデータを用意
- ② BERTopicを用いてトピックモデリングを行う
- ③ 分類されたデータのみで矛盾検知を行う（元論文の手法も実施）



データセット

テストデータ作成

- ① 5つのテーマを決めて、要求をグループ分けする
- ② 矛盾を均等に仕掛ける（各グループ3個ランダム抽出）
- ③ クラスタリング後LLM手法用のテスト（1グループあたり平均12.6個の要求）
- ④ 15組の矛盾正解ペアリスト保存
- ⑤ プロンプト指示手法用の元の要求と追加した矛盾を一つにした巨大な文書を保存



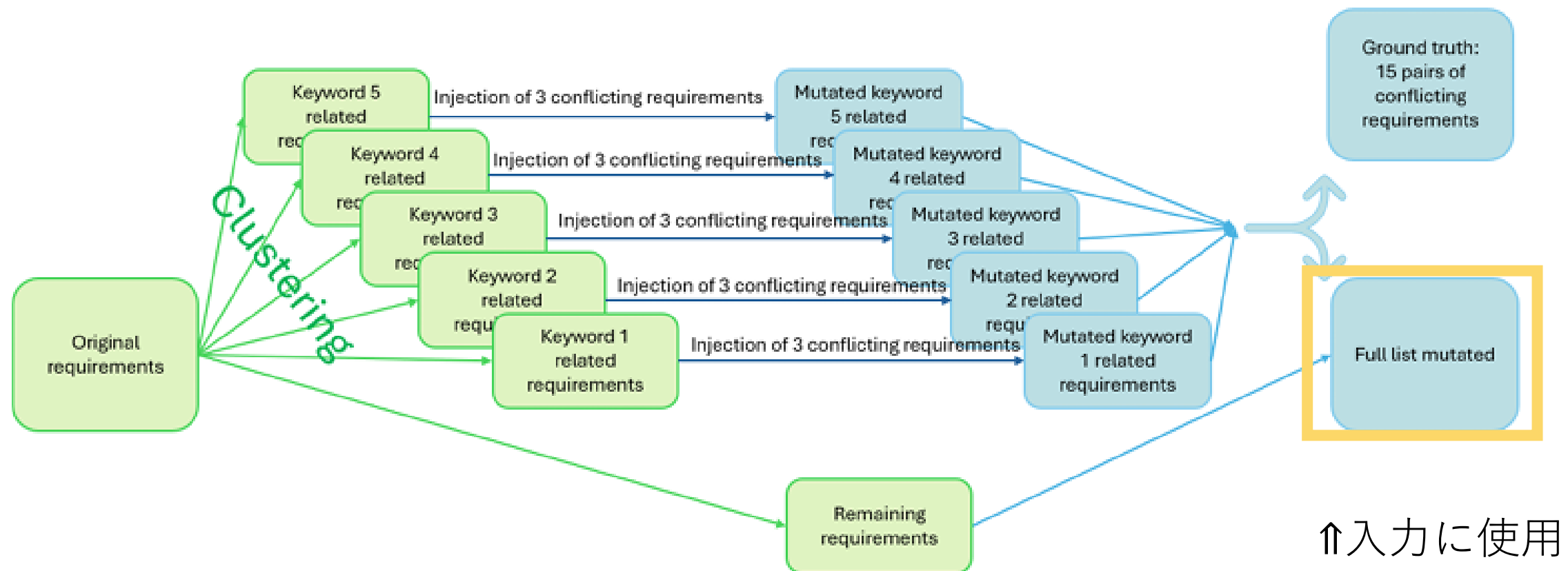
「1」

Fig. 1. Data Preparation

BERTopic

BERTopicにデータセットを入力する

- ・ クラスタ数や分類の強度については何も指定せず入力
⇒ 4 クラスタとその他の集めた1つのグループを作成

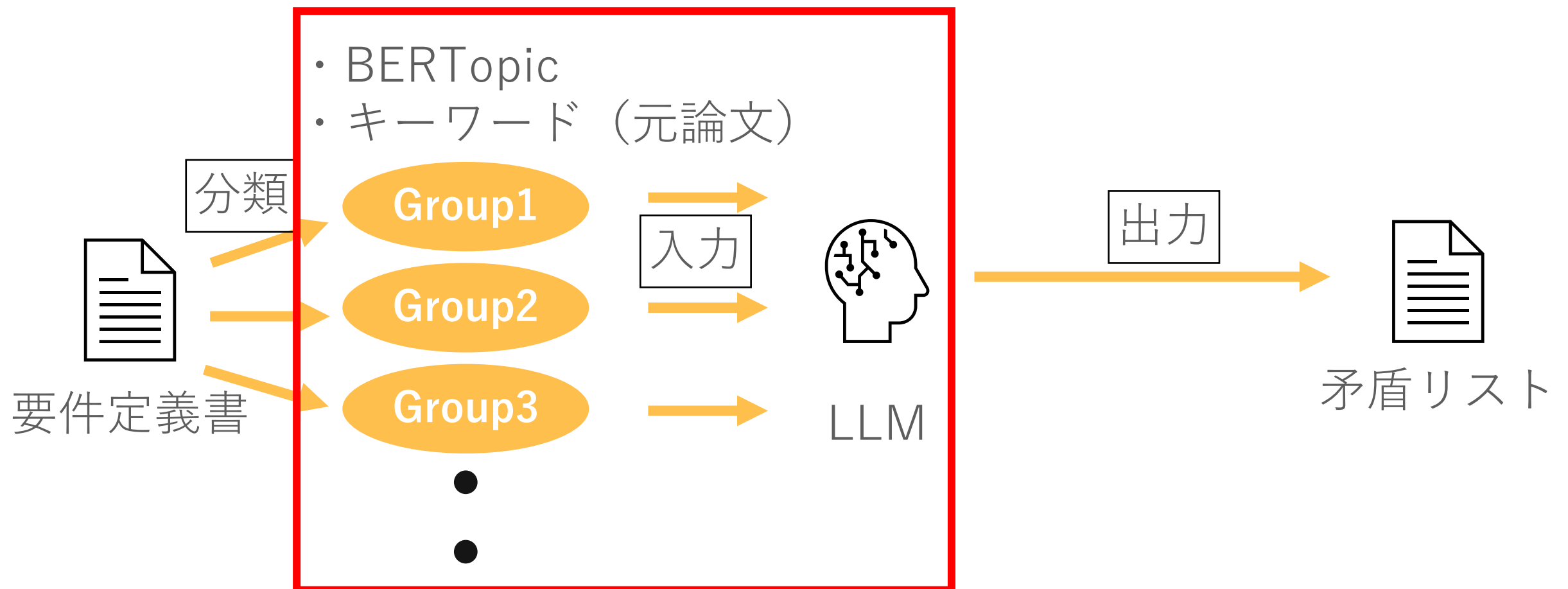


「1」 Fig. 1. Data Preparation

LLMによる矛盾検知

分類されたデータをグループごとにLLMに入力

- 提案手法：BERTopicによる分類
- 先行研究：キーワードによる分類
- 分類を行わない全件入力



出力された矛盾ペア

| 正解ラベル付け

◎意図的に仕込んだ矛盾

○仕込んだ矛盾によって発生した意図しない矛盾

○元々存在していた未発見だった矛盾

×矛盾していない

? 元論文でも出てこなかった要件のペア

評価用チェックシート

4つの判定区分

- ・ 明確な矛盾
- ・ 条件付きで矛盾
- ・ 矛盾なし
- ・ 情報不足

判定軸

4つの矛盾の類型

- ・ **時系列、実行タイミングのねじれ**

ー処理の「前後関係」や「前提条件」が入れ替わっている

- ・ **対象、処理範囲のずれ**

ーデータ構造の変更範囲や、特定の操作が影響を及ぼす「対象の定義」が衝突している

- ・ **認証主体、実行場所の違い**

ーシステムが「どこで」「誰が」処理を承認するかが異なる

- ・ **データフロー、制御権限の不一致**

ーデータの流れる経路や、そのデータを「誰が管理・閲覧・転送してよいか」のルールが矛盾している

判定の例

判定

- ①通信障害時や電波の届かない店舗であっても決済を行えるよう、**オフライン状態**での決済処理を許可し,完了できなければならない.
- ②残高の改ざんや二重支払いを確実に防止するため,すべての決済処理において必ず**オンライン通信**を行い,サーバー側の承認を得る必要がある.

判定区分：明確な矛盾

矛盾の類型：認証主体、実行場所の違い（どこで：オンライン・オフライン、誰が：端末側・サーバー側）

今後の予定

元論文でも出てこなかった要件のペア

- ・ 複数名に判定をしてもらう

⇒ エクセルで評価シートの作成

精度を評価

- ・ 評価指標を設定
- ・ 分類無し、キーワード分類、BERTopic分類のそれぞれの精度を評価