

自動分類におけるデータセット作成時の ミスラベルの調査：非機能要件分類を例として

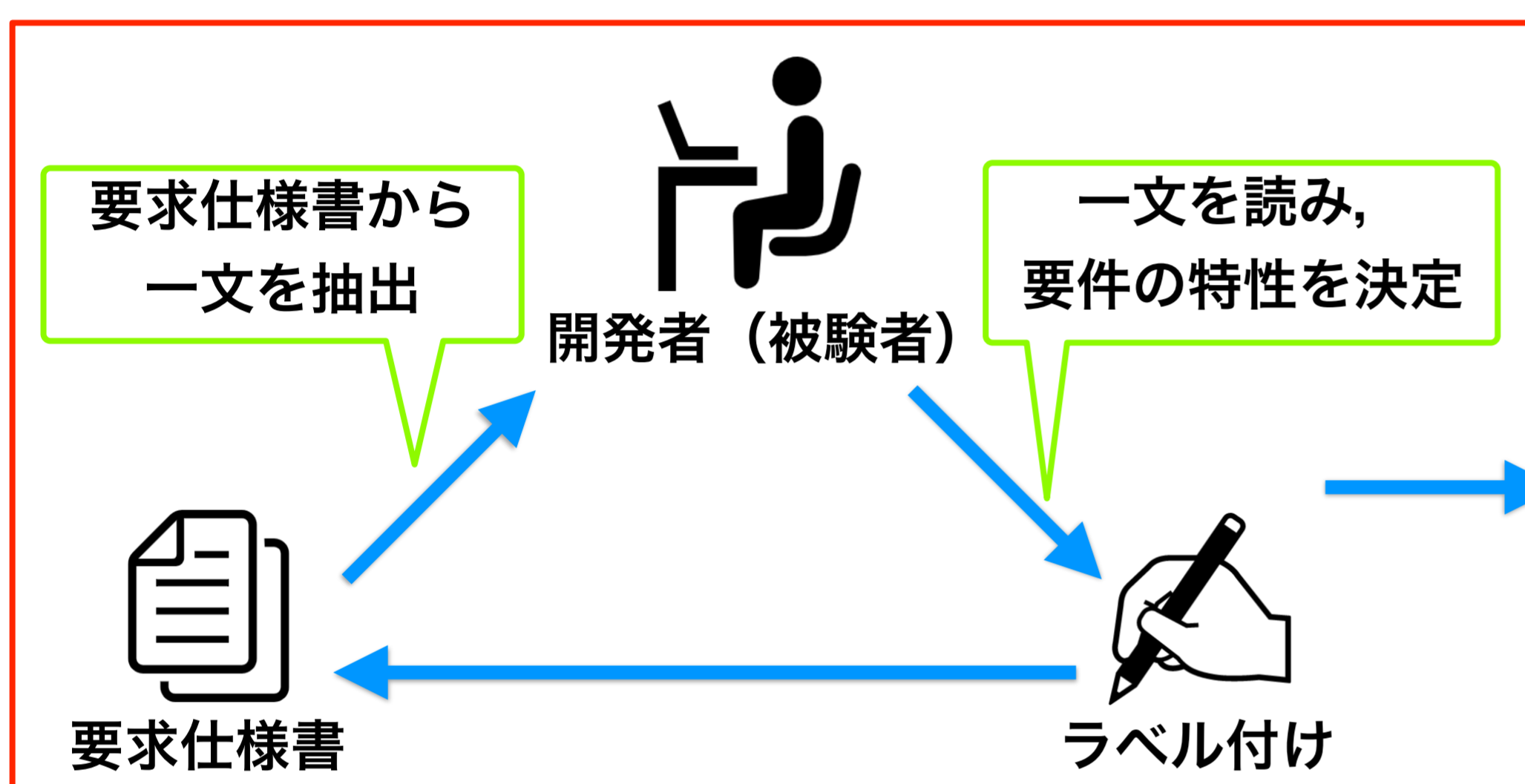
和歌山大学 木村 祐太 大平 雅雄

背景

ソフトウェア開発では、利用者のニーズを明確にする要求分析が行われ、その要件は大きく2つに分けられる。

- ・機能要件 : システムに要求される機能
- ・非機能要件 : システムの品質特性

要件は、設計・実装・テストで用いられるが、要件の把握、特に**非機能要件の把握は困難**。

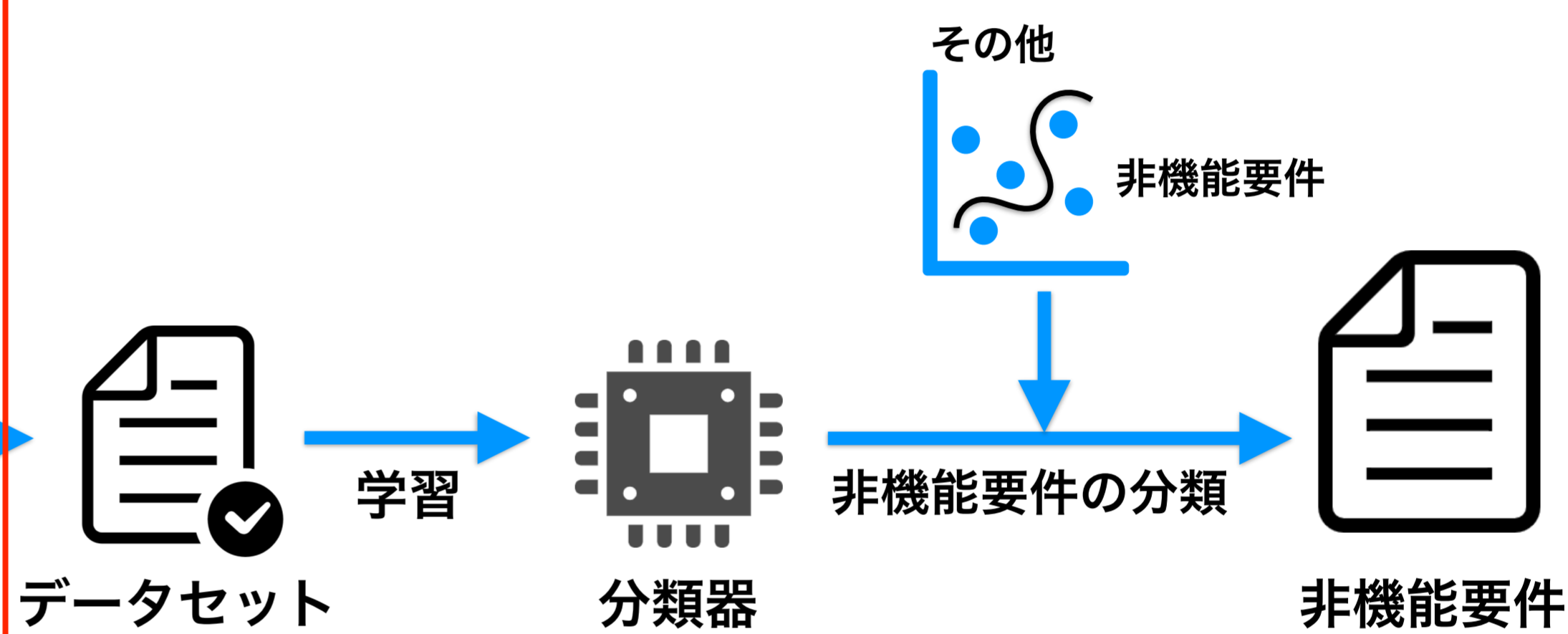


従来研究の非機能要件自動分類手法

動機と目的

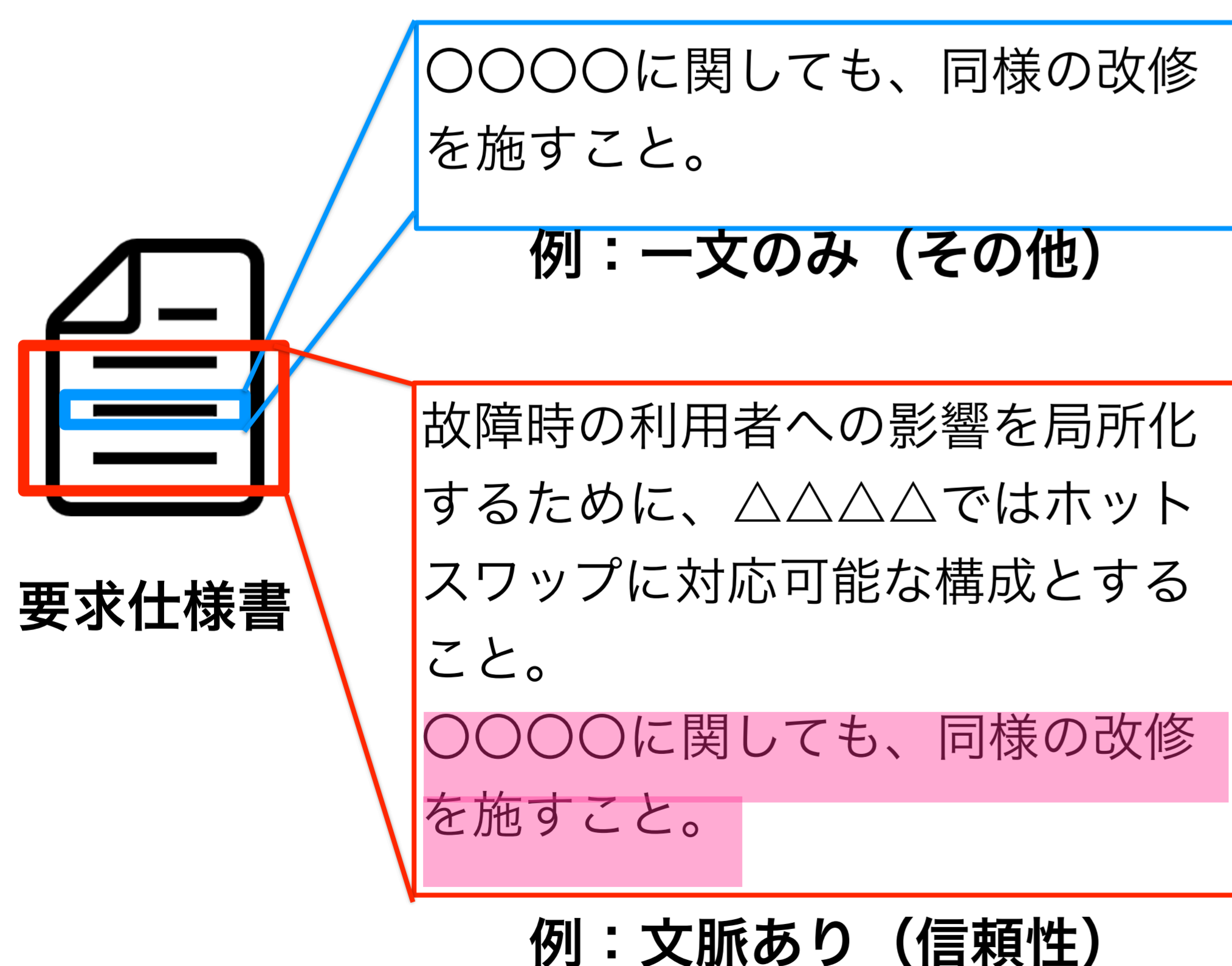
従来研究における非機能要件自動分類手法で用いられるデータセットは、**要求仕様書から抽出した要件（一文）にラベル付けを行うため、文脈は考慮されない**（下図赤枠）。

一文のみのラベル付けと文脈を考慮したラベル付けの結果を比較し、ラベル付け結果にどのような影響があるのかを明らかにする。



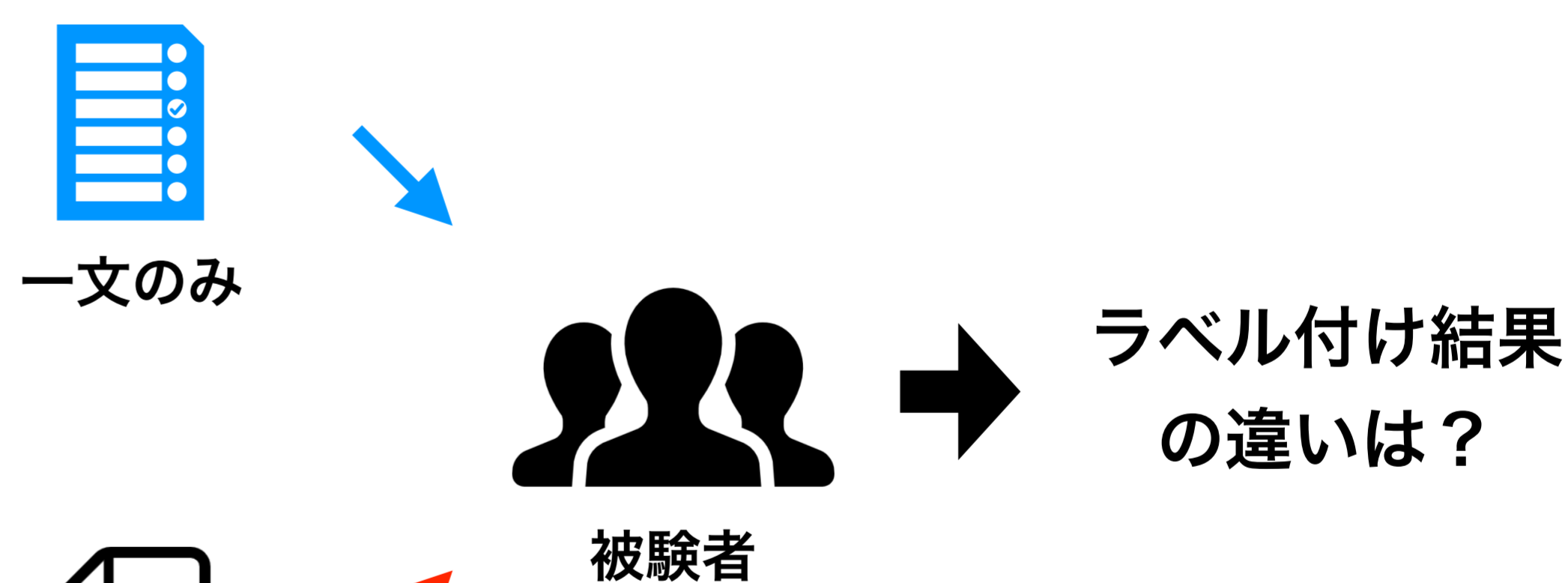
アプローチ

データセット



一文のみ、文脈ありのデータを用意しラベル付け*を行う。文脈ありは、ラベル付け対象の要件に色付けを行い、わかりやすくしておく。

被験者実験（ラベル付け）



ラベル付け結果の違いを調査するために被験者実験を行う。

学習効果のカウンターバランスを取るために、

- ・一文のみ → 文脈あり
- ・文脈あり → 一文のみ

2通りで実験し、ラベル付け結果に違いがあるのか統計的検定を用いて検証する。

* ISO/IEC 25010: <https://iso25000.com/index.php/en/iso25000-standards/iso-25010>