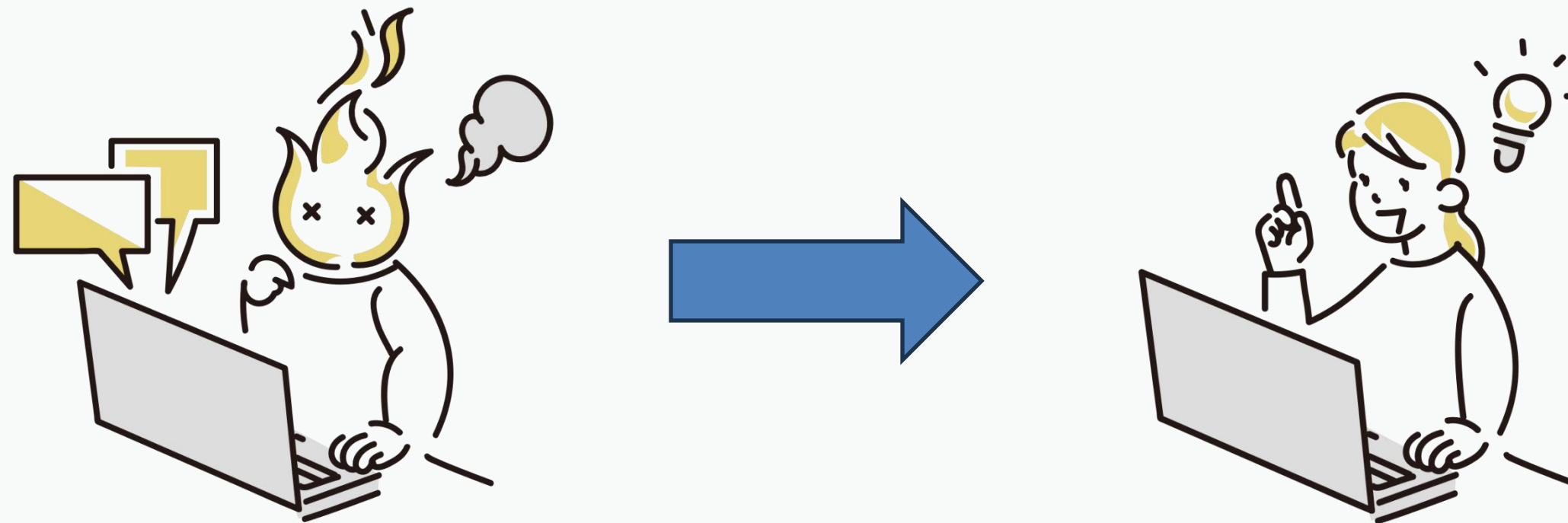


# 明日から不満が見たくなる

## -不満マイニング-



# 不満①

どんな歌を歌っても響かないね

音痴。歌手やめればいいのか。

どこが良いんだかわからない



※「不満調査データ」より、個人や組織の名前を除き一部改変したものを記載

## 不満②

大きい会場は席によって  
価格を変えて欲しい

コンサートグッズを  
ネット通販でも売ってほしい

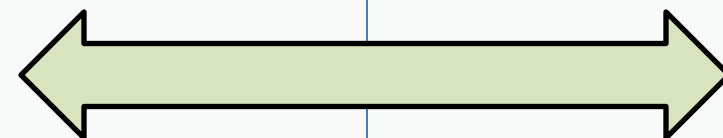
大体でいいから終演時間を  
あらかじめ教えて欲しい



※「不満調査データ」より、個人や組織の名前を除き一部改変したものを記載



何が違う？



どんな歌を歌っても響かないね

音痴。歌手やめればいいのに。

どこが良いんだかわからない

大きい会場は席によって  
価格を変えて欲しい

コンサートグッズを  
ネット通販でも売ってほしい

大体でいいから終演時間を  
あらかじめ教えて欲しい



個人的な感情・感想

どんな歌を歌っても響かないね

音痴。歌手やめればいいのに。

どこが良いんだかわからない



具体的な要望・提案

大きい会場は席によって  
価格を変えて欲しい

コンサートグッズを  
ネット通販でも売ってほしい

大体でいいから終演時間を  
あらかじめ教えて欲しい

人もLLMも「役に立つとは言えない」と判断した不満

どんな歌を歌っても響かないね

音痴。歌手やめればいいのに。

どこが良いんだかわからない

人もLLMも「役に立つ」と判断した不満

大きい会場は席によって  
価格を変えて欲しい

コンサートグッズを  
ネット通販でも売ってほしい

大体でいいから終演時間を  
あらかじめ教えて欲しい

## 本研究の背景

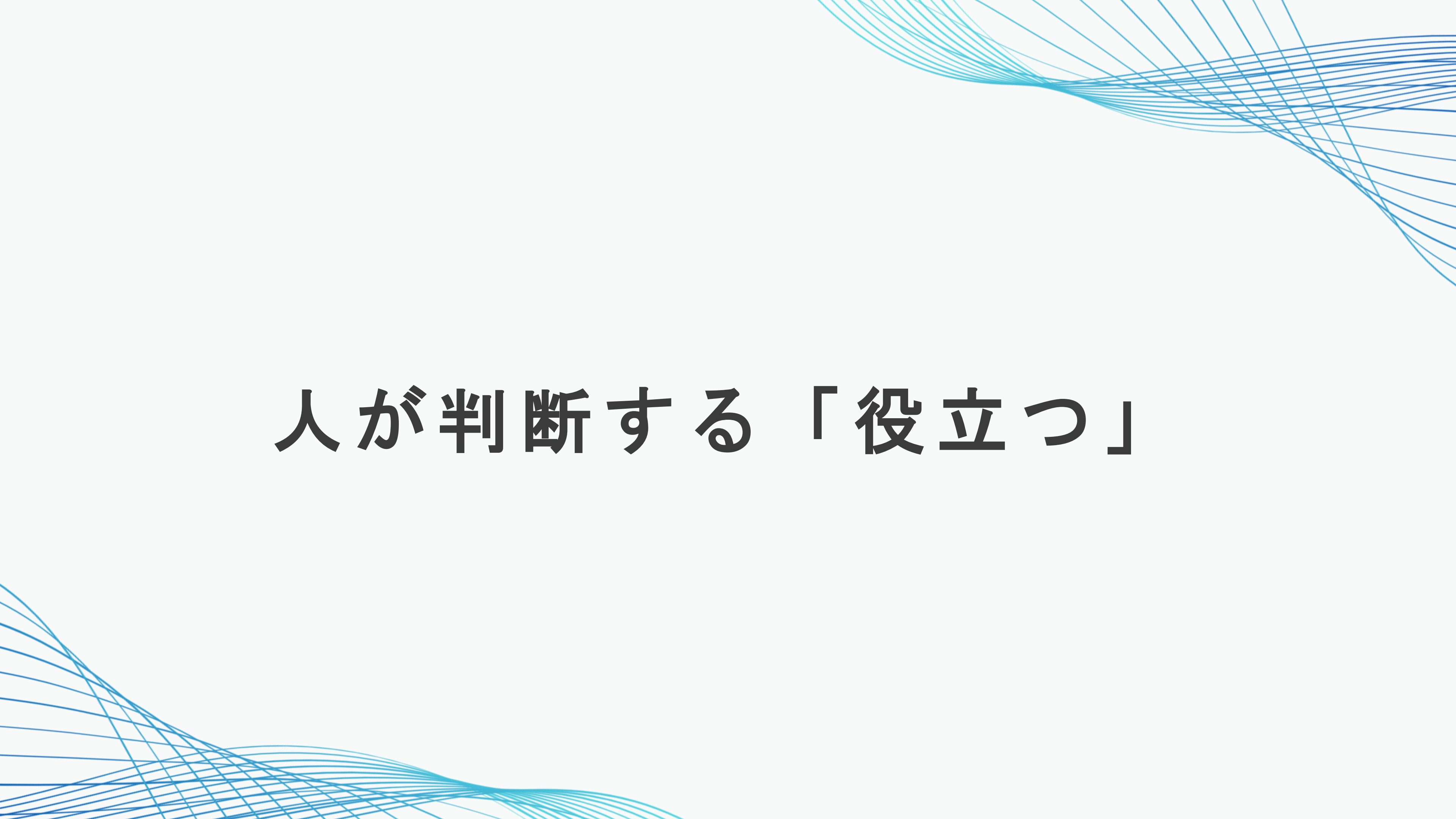
最終目的：不満を有効活用することで世の中の不満を減らしたい！

➡ 不満を整理することで有効活用できないか？

### 不満を整理することのメリット

- コメントからコンテンツの問題点を洗い出すことができる
  - どれだけの人が同じことを思っているのかもわかる
- 見ていてただ不快になるコメントの表示優先順位を下げ、知りたい情報を優先的に閲覧することができる

不満を整理するためにまずは「**見たくなる不満**」を探そう



人が判断する「役立つ」



## 前回の研究

### 「役立つ不満」抽出のためのコーパス構築とトピック分類

#### 不満

- 不満自体が**悪いものではない**
- 不満を表に出すことで物事の解決のきっかけになることもある

#### 役立つ

- “役立つ”という言葉は**曖昧**である
- 言葉ひとつでも何に向けての発言かによって変わってくる

【例】「散らかってるから何とかして」

親のセリフ？ イベント会場のゴミ箱周りについて？

⇒ **トピック** を分けることで判断しやすく

# 扱うデータをトピックで絞る

- 対象とするデータ
  - 不満調査データのサブカテゴリ「ライブ・コンサート・カラオケ」
  - 17,529件

株式会社Insight Tech(2017), “不満調査データ. 国立情報学研究所情報学研究データリポジトリ.(データセット),” 2016.  
<https://doi.org/10.32130/idr.7.1>

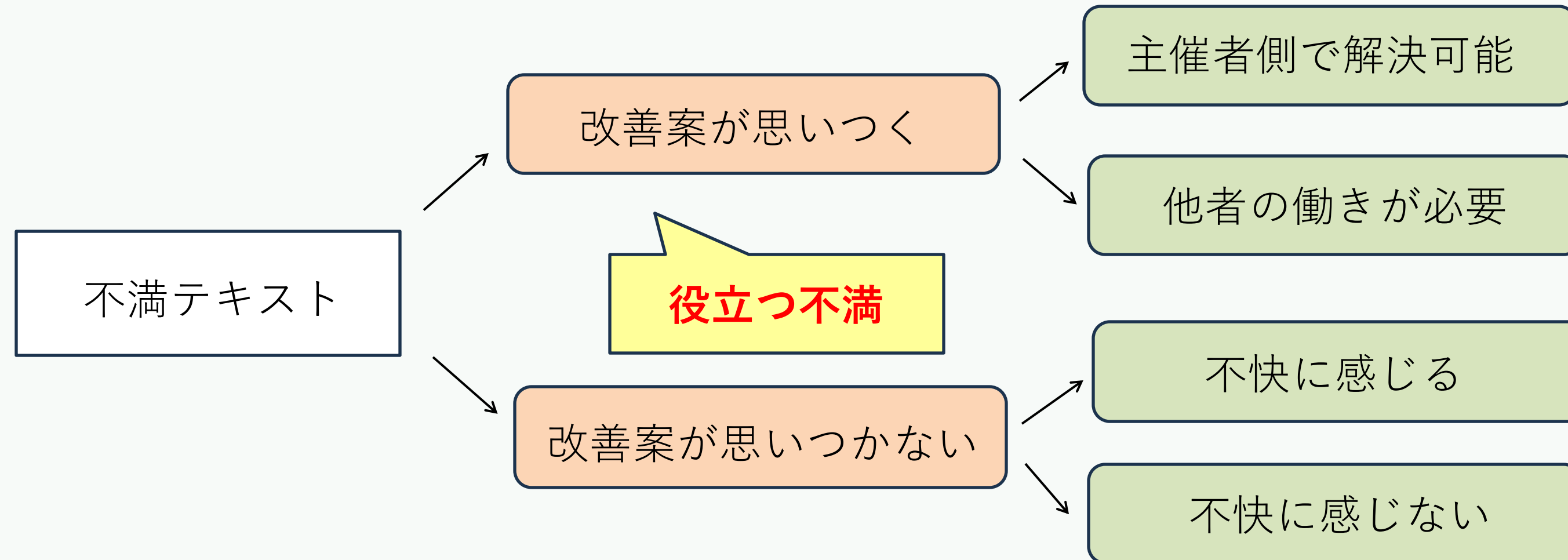
トピック分類をして  さらにトピックを絞る

- ライブのグッズ
- ライブの座席やチケット代
- チケットが取れない
- カラオケの部屋や採点
- カラオケのフリータイム

妥当な件数のサンプルを抽出したうえでアノテーションする  
今回は1,612件

## 「役立つ不満」の評価基準を設定

- 「改善案が思いつく = 役立つ不満である」と定義し、ラベルを設定



- 3名のアナテーターに直感的に回答してもらった
- 評価がわかれた場合には、多数決でラベルを決定した

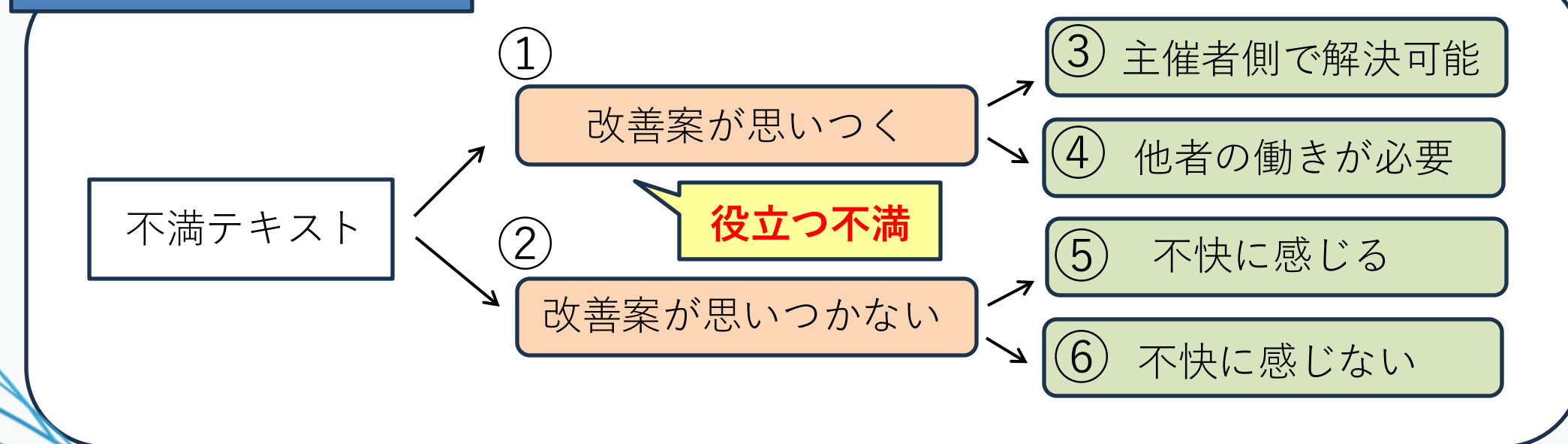
# アノテーションの結果ーサブデータセットの作成ー

## 【サブデータセットの一例】

1人分のラベル

text	ans1-1	ans1-2	ans2-1	ans2-2	ans3-1	ans3-2
コンサートで後ろの席でも前の席と同じ値段なことが不満です。	1	3	1	3	1	3
ファンクラブに入らなくてもコンサートのチケットを買えるようにしてほしい。	2	6	1	3	1	3
カラオケで隣の人の声が聞こえてて自分が歌いにくいです。	1	3	1	3	2	6

### ラベルの種類



- トピックごとのアノテーション結果も閲覧可能
- 多数決でのラベル結果も記載
- 有効回答1,601件のデータ数

ある程度の基準を設けても人の**性格**や**価値観**によって**評価は変わる**

## 昨年いただいた意見と課題

### 良かったポイント

- 不満というネガティブなものをポジティブに捉えているところが良い
- より良いデータセットを考案しているところが良い
- 不満を役立てようとしているところが良い

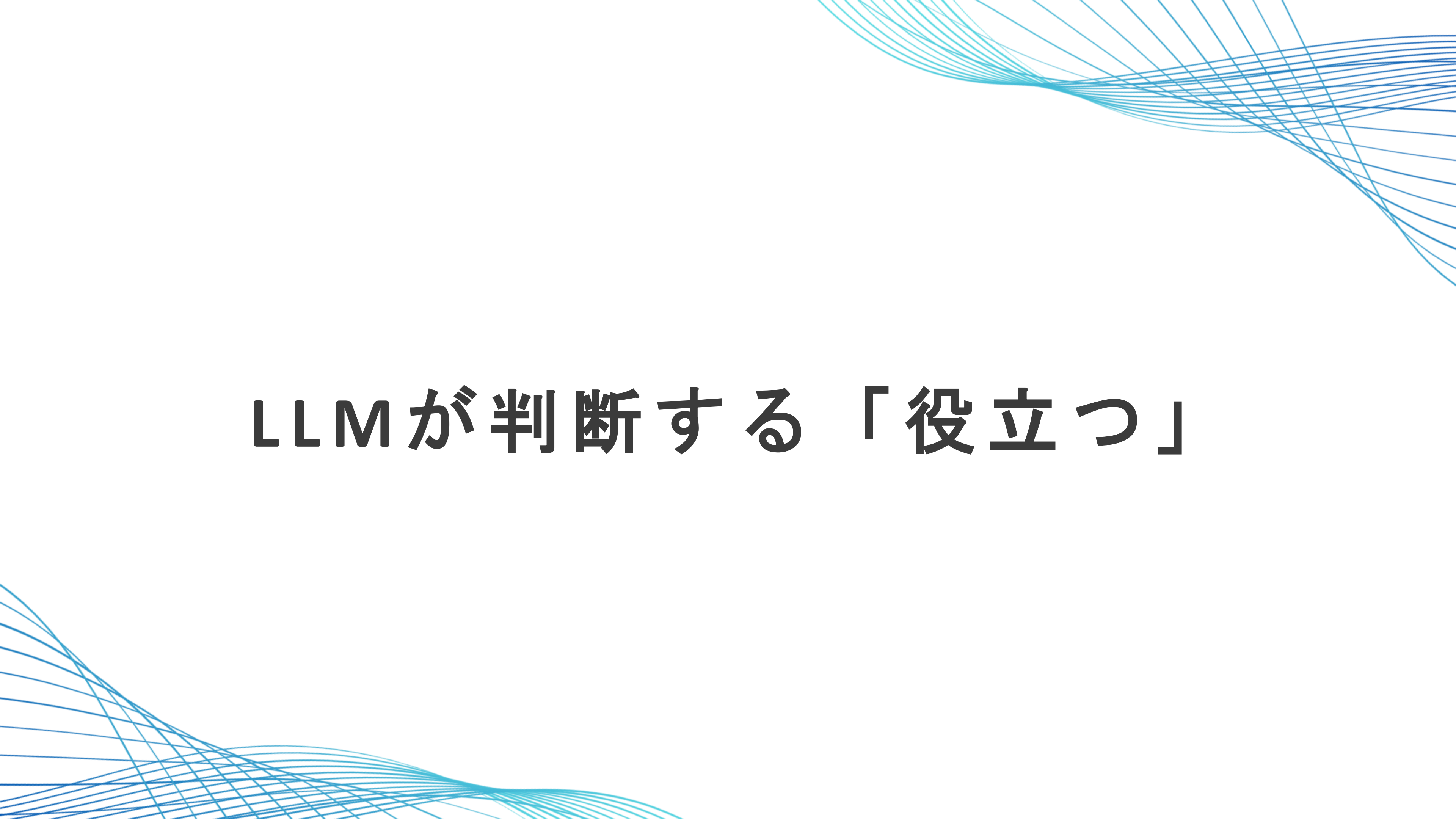


### ありがたいコメント

- ラベルはより客観的指標にしないと価値観の違いが出てしまう
- 不快に感じるかどうかは独立しているのでは？
- 改善案が思いつくかどうかはその人の経験に基づくのでは？



**人によって価値観が違うことや、評価基準の選定が課題！**



**LLMが判断する「役立つ」**

## LLMを使うことの意義

### 現状の課題

一般的な人の価値観を知る方法のひとつとして  
たくさんの人にアンケートをとる必要がある

発想の転換をすると…

- LLMはさまざまな考え方を学習させることによって一般的な回答を出力するものである

⇒ 代わりにLLMに聞くことで一般的な考え方に近い回答が得られると考えられる

### LLMを使う上での課題

LLMによって学習しているデータは違う  
一般的な考え方を知るには複数のLLMの結果を得る必要がある

## 今回の研究

# LLMが考える「役立つ不満」とは

前回の研究を経ての疑問点

LLMは何を基準に「役立つ」とするのだろうか

- 人は基準がないと答えようがないが、LLMは無理やりでも答えようとする

⇒ LLMの回答を見ることで「役立つ」基準が見えるかも

不満の文章は※「発話の役割」に特徴があるのか

- 「発話の役割」という新たな指標で不満を分類できないか

⇒ LLMに発話の役割を判断してもらうことで不満の特徴が見えるかも





## アプローチの詳細

- 対象とするデータ
  - 不満調査データのサブデータセット有効回答1,601件のデータ
- 使用するLLM
  - Llama3-ELYZA-JP-8B
  - Llama3-Swallow-8B-Instruct-v01
  - Qwen2-7B-Instruct

前回の研究で作成したもの

発話の役割38種

事実	補足	価値判断
知識取得	知識提供	教える
教えられる	依頼	確認
行動要求	発言要求	提案
礼	謝罪	賛成
反対	行動要求受入	行動要求受入納得
理由陳述	保持	納得
話題転換	遠慮	平静
命令	反省	笑い
戸惑い	反発	挑発
尊敬	嫉妬	嫌味
非難	好意	肯定
否定	その他	

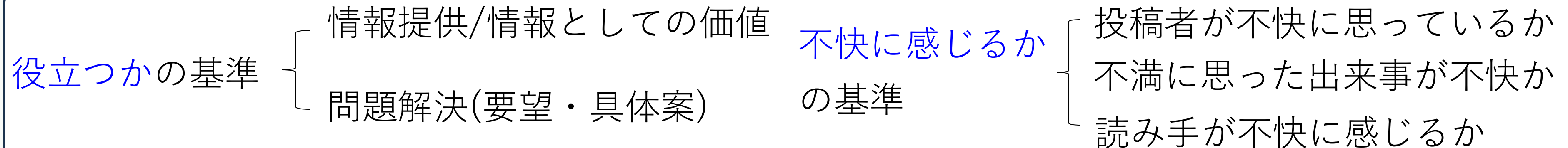
3つを比較することで基準としている**共通点**を調べる

### LLMに尋ねる内容

1. ラベルの選択(「役に立つ」 or 「役に立つとは言えないか」)
2. ラベルを選択した理由
3. 発話の役割の中から最も適する1つを選択
4. 不快に感じるかどうか(はい/いいえ)
5. 不快に感じるかどうかの理由

## 考察結果

### LLMが考える評価基準



### LLMが考える発話の役割

知識獲得・知識提供・価値判断・行動要求 など

LLMは**不満＝純粋な情報**として捉えており  
 テキストに書かれた内容が**情報として価値があるか**を判断している

求める評価基準を明記することで欲しい「役立つ不満」を閲覧できる可能性がある



# 人とLLMの比較

## 人とLLMの比較

### ■ 人、LLM = 「役立つ」

※人は3名のラベル付け結果が一致しているものを表示  
 ※LLMは3つのラベル付け結果が一致しているものを表示

大きい会場は席によって価格を変えてほしいです。

LLMの理由：価格の差異により、会場の座席の需要と供給のバランスを取ることができ、効率的で公正な価格設定が実現できるため。

受験生が気の毒。試験の日に重ならないように配慮してあげればいいのに

LLMの理由：この発話は、受験生のストレスや困難を理解し共感し、具体的な行動を提案しているため。

### ■ 人、LLM 「役に立つとは言えない」

歌聞いたこと無いけど、最低。聞く気失せたわ。

LLMの理由：この発話は、歌を聞いたことがない人に対して「最低」というネガティブな評価を下し、聞く気を失せたと発言しているため。

音痴がなおらないから嫌になります

LLMの理由：音痴が治らないこと自体は、医学的、心理学的、教育的観点から見て、解決策や治療法が確立されていないため。

## 人とLLMの比較(続き)

### ■ 人 = 「役立つ」、LLM = 「役立つとは言えない」

※人は3名のラベル付け結果が一致しているものを表示  
※LLMは3つのラベル付け結果が一致しているものを表示

チケット購入したら偽造だった。

LLMの理由：偽造チケットが手に入ってしまったため、使用することができず、役に立たないから。

コンサートのチケット代が前でも後ろでも同じ値段なのはズルい。

LLMの理由：チケット代が前でも後ろでも同じ値段なのは、客の視点や体験を考慮していないから。

### ■ 人 = 「役に立つとは言えない」、LLM = 「役立つ」

第1日目が土曜日だが、土曜の午前は授業があって行けない。

LLMの理由：土曜の午前には授業があるため、日程の都合がつかないという具体的な情報が提供されているため。

すごくいい曲で、大好きなのですが、頭から離れませーん!!

LLMの理由：この発言は、楽曲についての感想を述べているため、「役に立つ」と判断しました。

補足：LLM間でも出力結果に差はあった

## 今後の展望

- 「役立つ不満」の **カスタマイズ**
- それぞれの人が考える役立つ不満の評価基準を設定することで **価値観に合わせた** 「役立つ不満」を表示させる
- 事実、状況報告、提案 etc...

どんな不満をお探しですか？

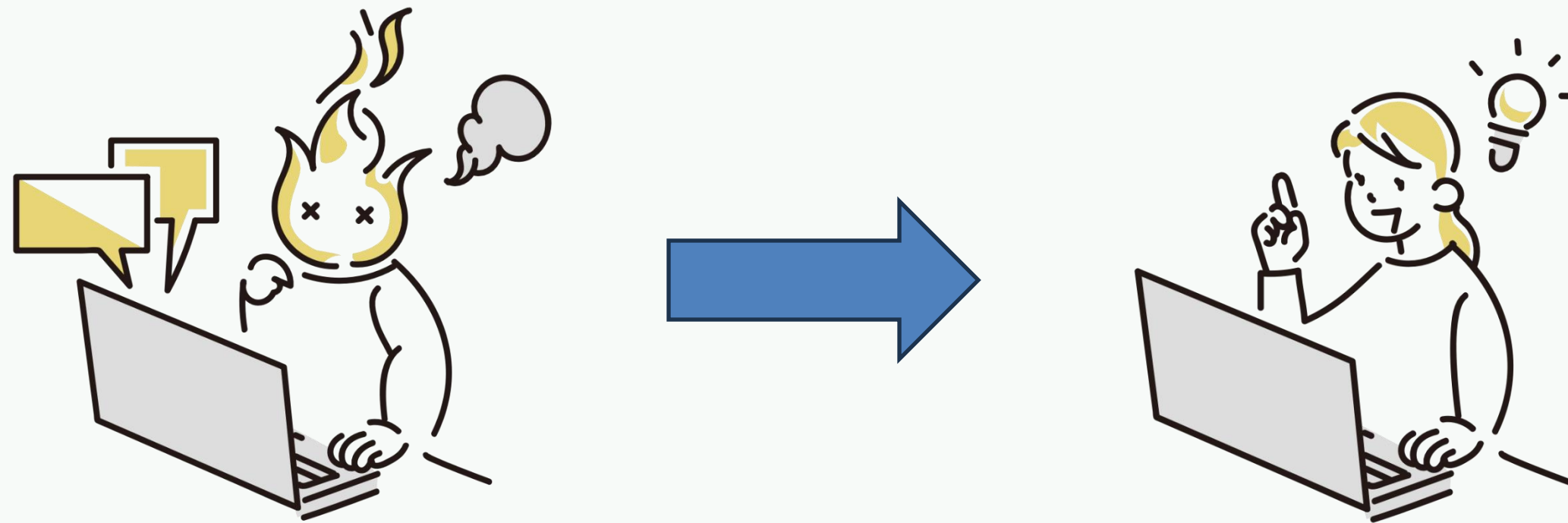
  

- 特定の基準を設けて **不満を分類**
- **既存の指標を軸に** 不満を分類できるかを分析する
- 発話の役割、利用規約 etc...

## タイトルの振り返り

明日から不満が見たくなる

-不満マイニング-



# まとめ

最終目的：不満を有効活用することで世の中の不満を減らしたい！

## 前回の研究

- 「改善案が思いつく = 役立つ不満」と定義して人によるアノテーションを実施
- 使用する不満をトピックに分類することでより評価しやすく設定した
- サブデータセットを作成し、ある程度評価基準を絞っても人の性格や価値観によって回答が変わることがわかった

## 今回の研究

- LLMが「役立つ」をどういう基準で評価するかを分析
- 3つのLLMを使用することで評価基準の共通点を調べた
- LLMに回答させることで複数の「役立つ」という考え方が明らかになった

## 今後の展望

役立つ不満のカスタマイズや、評価軸を設定した上での不満の整理