



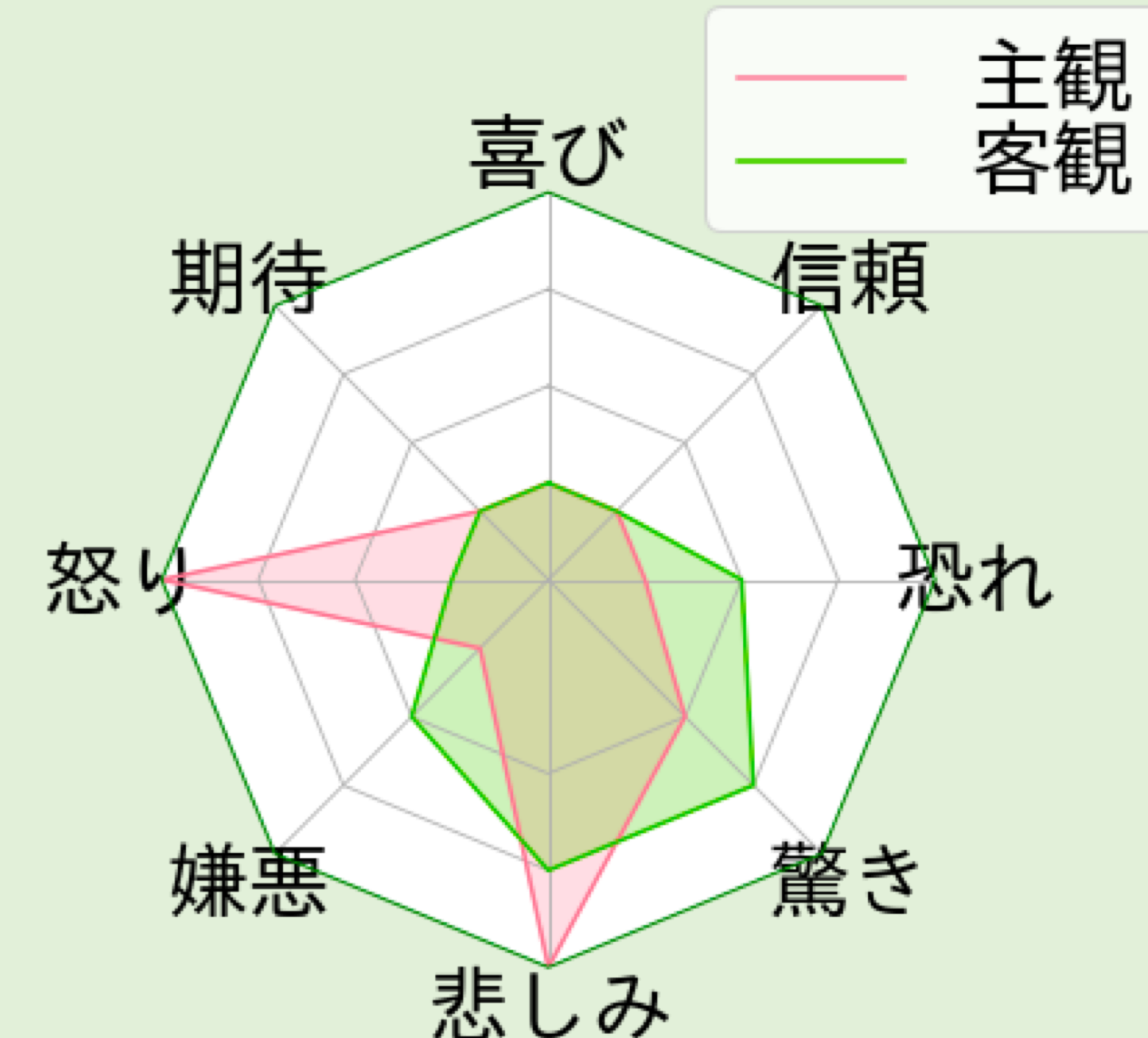
## 日本語の感情分析のためのデータセットを構築

<https://github.com/ids-cv/wrime>

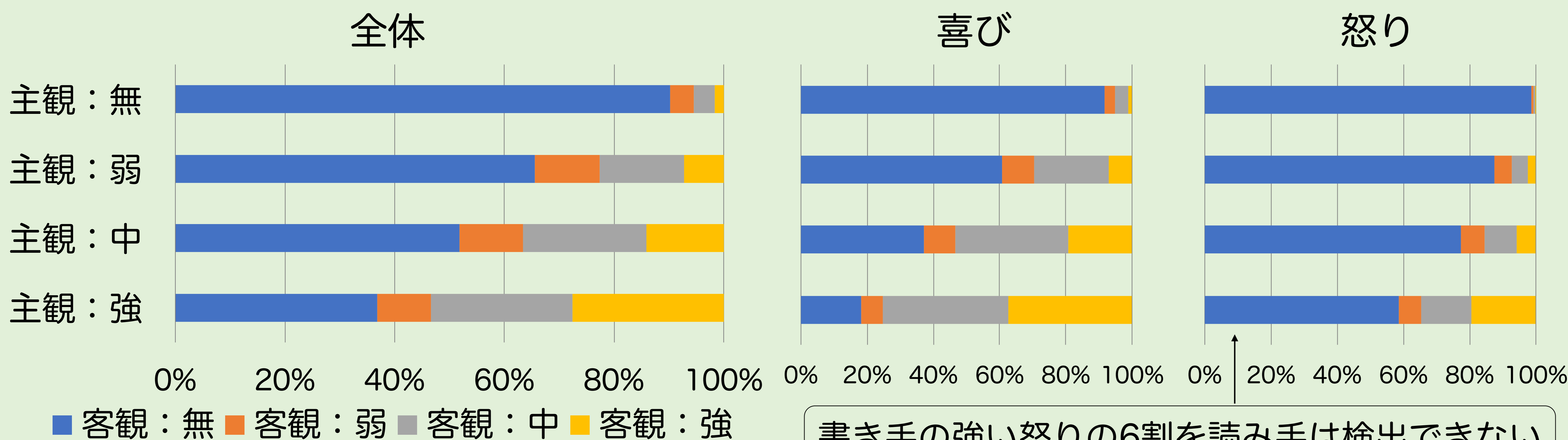
- 50人の書き手による合計 17,000件の日本語SNSテキスト
- Plutchikの8感情の強度を主観（書き手）と客観（読み手）の両方で4段階で付与  
※ 既存データは基本的に客観のみ
- 公開版は80人からの43,200件

タイヤがパンクしてた。。いたずらの可能性が高いんだって。。

	喜び	悲しみ	期待	驚き	怒り	恐れ	嫌悪	信頼
主観	0	3	0	1	3	0	0	0
客観A	0	3	0	3	1	2	1	0
客観B	0	2	0	2	0	0	0	0
客観C	0	2	0	2	0	1	1	0



## アノテーションの傾向：読み手は書き手の感情を過小評価しがち



## BERTによる感情強度推定

感情強度推定の平均絶対誤差

	主観感情	客観感情
最頻クラス	0.578	0.395
主観BERT	0.553	0.377
客観BERT	<b>0.519</b>	<b>0.317</b>

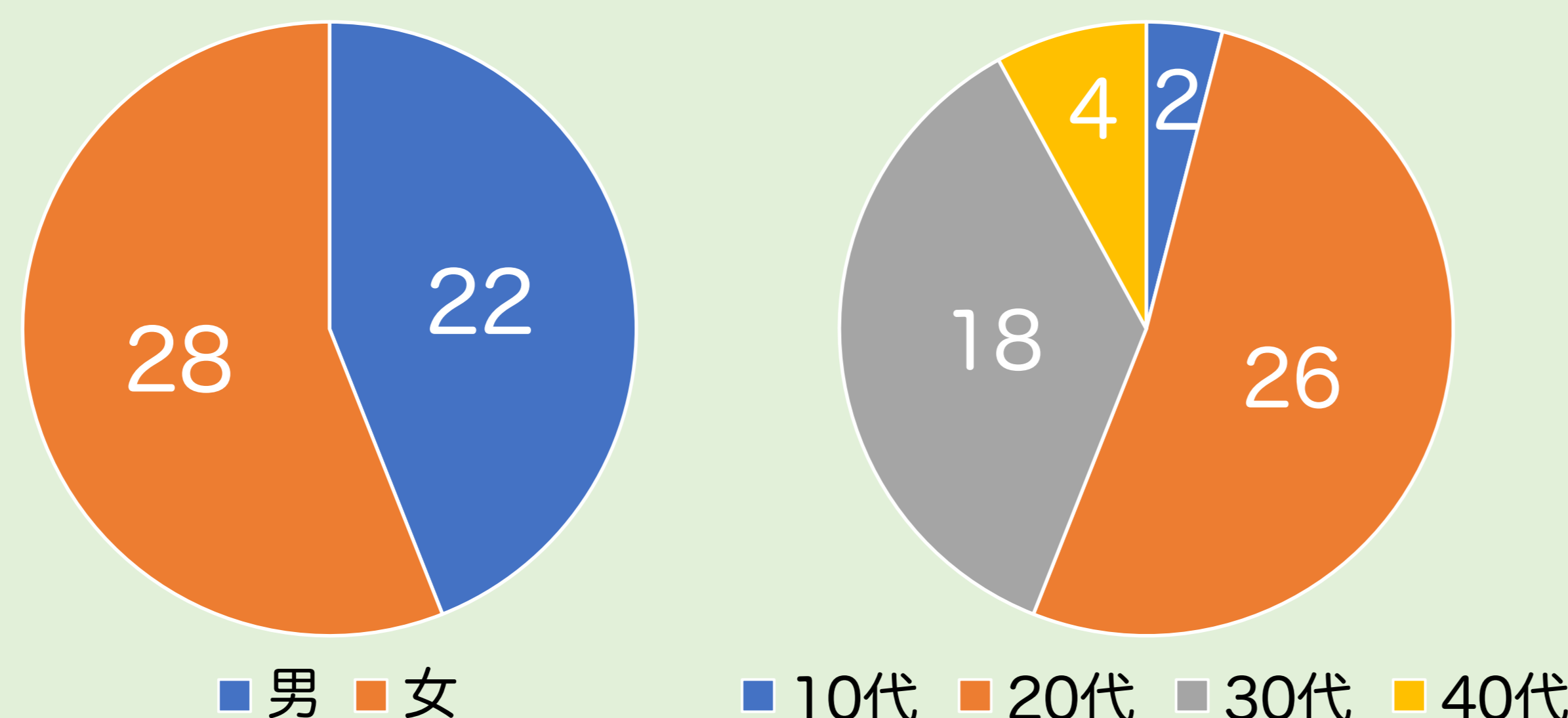
- 訓練：15,000 検証：1,000 評価：1,000
- 主観感情（書き手の感情）を推定する場合も客観データで学習した方が推定誤差が小さい

## アノテーションの詳細

※ 60項目の性格診断も実施（詳細は原稿をご覧ください）

### 主観ラベルの収集

- ランサーズで50人雇用
- SNS上の自身の過去の投稿に感情強度を付与
- 性別と年齢の内訳 →



### 客観ラベルの収集

- ランサーズで3人雇用
- 各投稿に（主観ラベルを見ずに）感情強度を付与
- 30代女性2 + 40代女性1

### 評価者間の一致率

	Kappa
客観 - 客観	0.55 ~ 0.59
主観 - 客観	0.44 ~ 0.47
主観 - 客観3人の平均	0.52