# Understanding Gender-related Discussions in Android Mobile Applications Through Reddit

Dariya Nagashibayeva

Supervised by Kuldar Taveter, PhD and Tahira Iqbal, MSc

UNIVERSITY OF TARTU
Institute of Computer Science
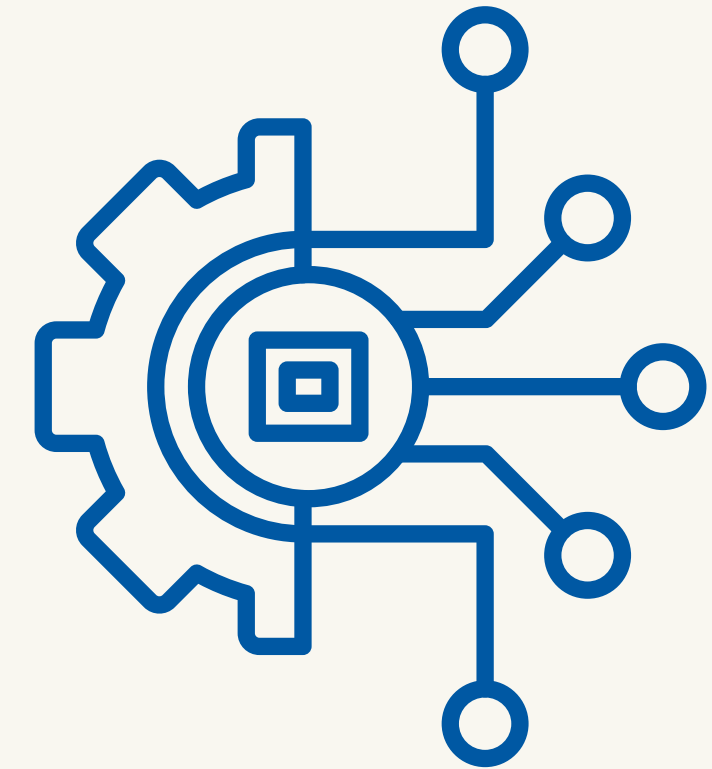1632

# Overview

# 01  Motivation

Gender-inclusive
mobile applications

Reddit as a source
for requirements

Automated
classification

# 02 Background

Previous work:

- M. Shahin et al. - A study of gender discussions in mobile apps [1]

- T. Iqbal et al. - Mining Reddit as a new source for software requirements [2]

- Humphrey O. Obie et al. - A First Look at Human Values-Violation in App Reviews [3]

- M. Fazzini et al. - Characterizing human aspects in reviews of COVID-19 apps [4]

# 03  Problem statement

- There is a lack of research on gender inclusiveness violations in mobile applications

- Current gender-related datasets are mostly derived from the same sources

# 04 Research questions

RQ1: Does Reddit data provide meaningful information about gender-related discussions in mobile applications?

RQ2: Can we further categorize the gender-related discussions in Reddit posts?

RQ3: How do different ML models used for the automated classification of Reddit data for gender discussions in mobile applications compare with each other?

# 05 Approach

## Initially defined steps:

- Data collection and quantitative analysis
- Data annotation and qualitative analysis
- Automated classification
- Categories of gender discussion
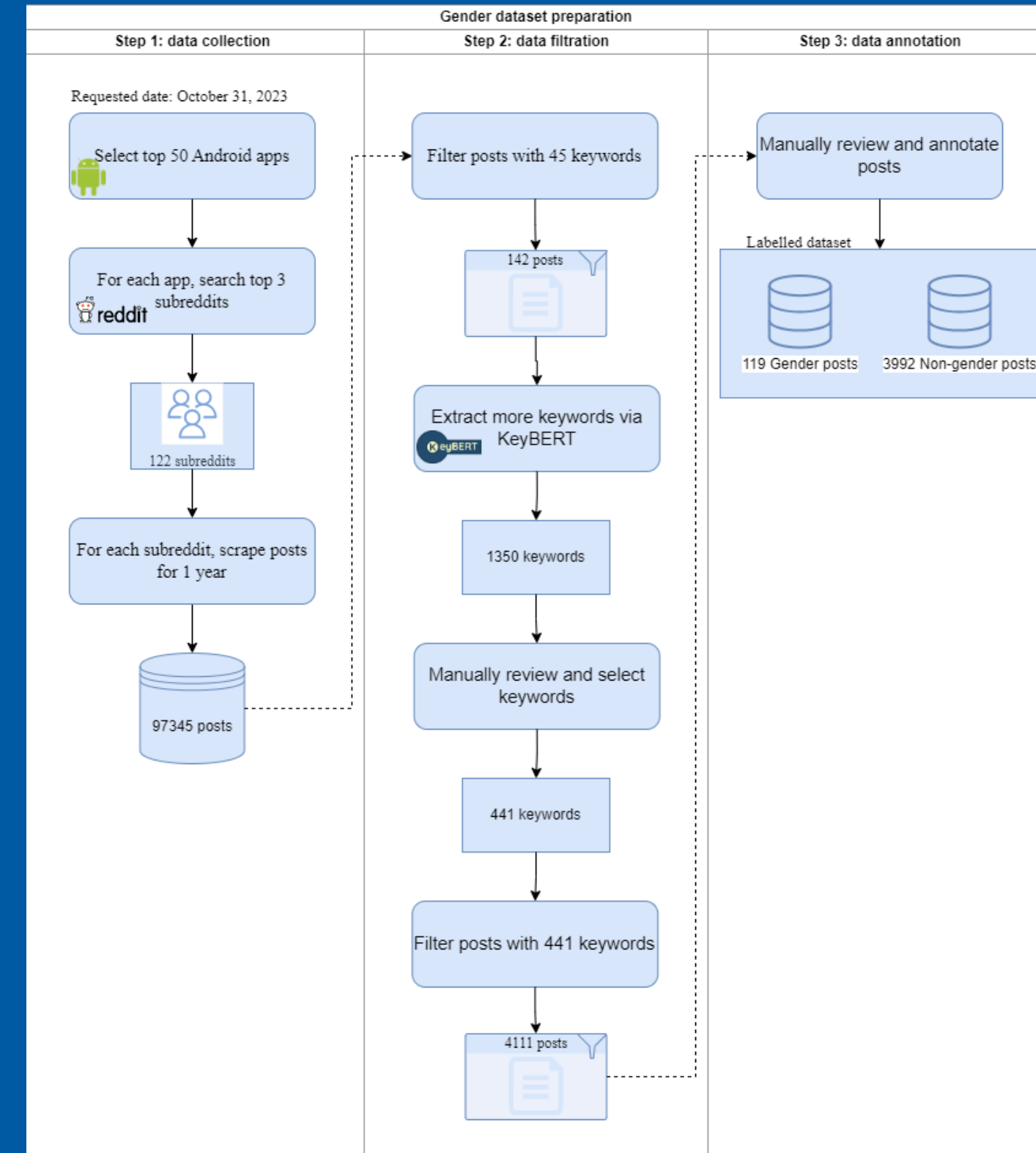
# 06 Results

## Selection of applications:

- Top 50 free Android apps
- Requested date: October 31, 2023
- Sources:
    - https://apptopia.com
    - https://play.google.com/store/apps

# 06 Results

Finding subreddits related to the applications:

- Top 3 subreddits associated with each app
- Criteria:
  - either official or directly dedicated to the app
  - at least 1000 members
  - open (available to download data)

# 06 Results

## Data collection:

- PRAW: The Python Reddit API Wrapper
- For each subreddit download posts for 1 year
- 97345 posts in total



Requested date: October 31, 2023

Select top 50 Android apps

For each app, search top 3 subreddits

122 subreddits

For each subreddit, scrape posts for 1 year

97345 posts
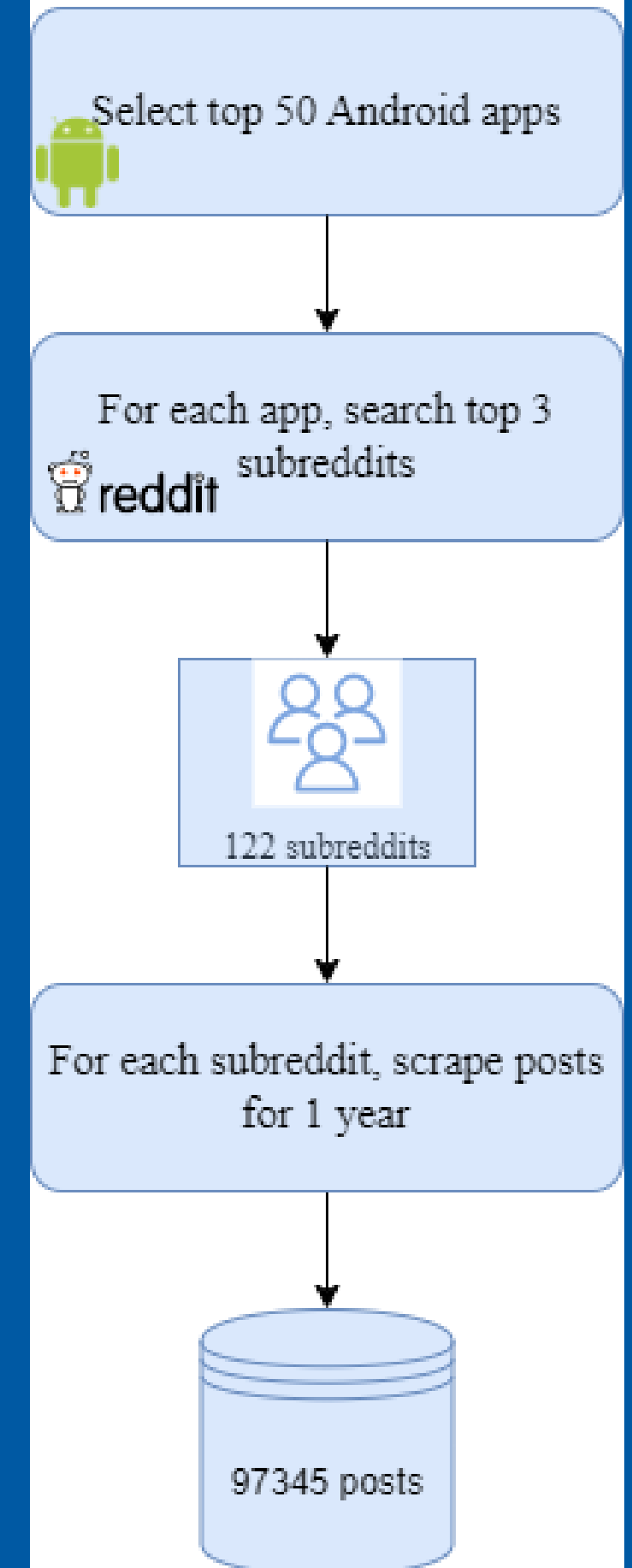
# 06 Results

## Keywords extraction and filtering:

- KeyBERT is used for gender-related keyword extraction
- 1350 keywords and key phrases extracted
- 441 keywords selected after manual review
- 4111 posts after filtering
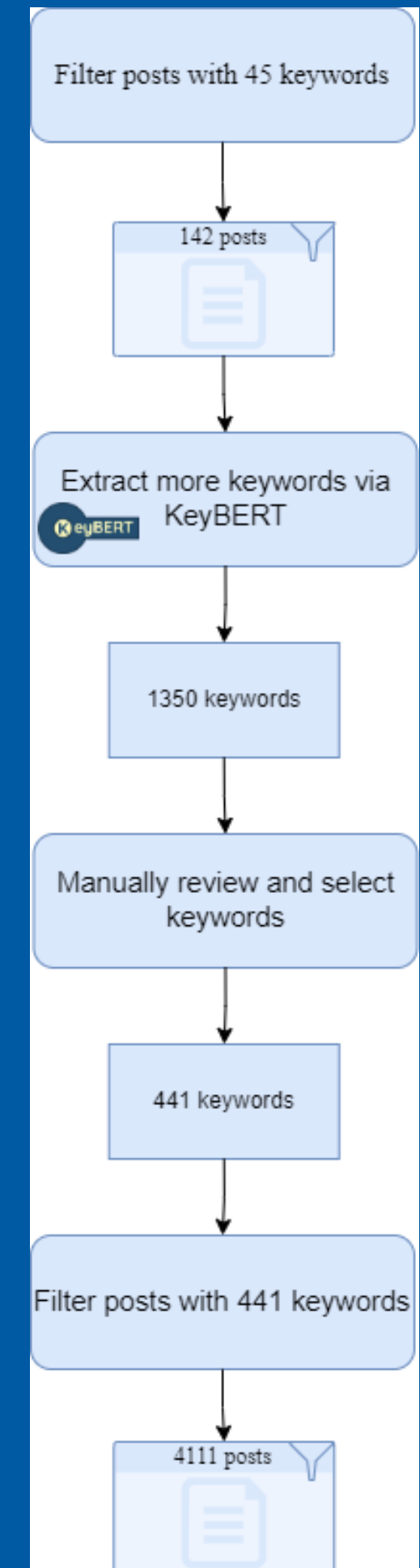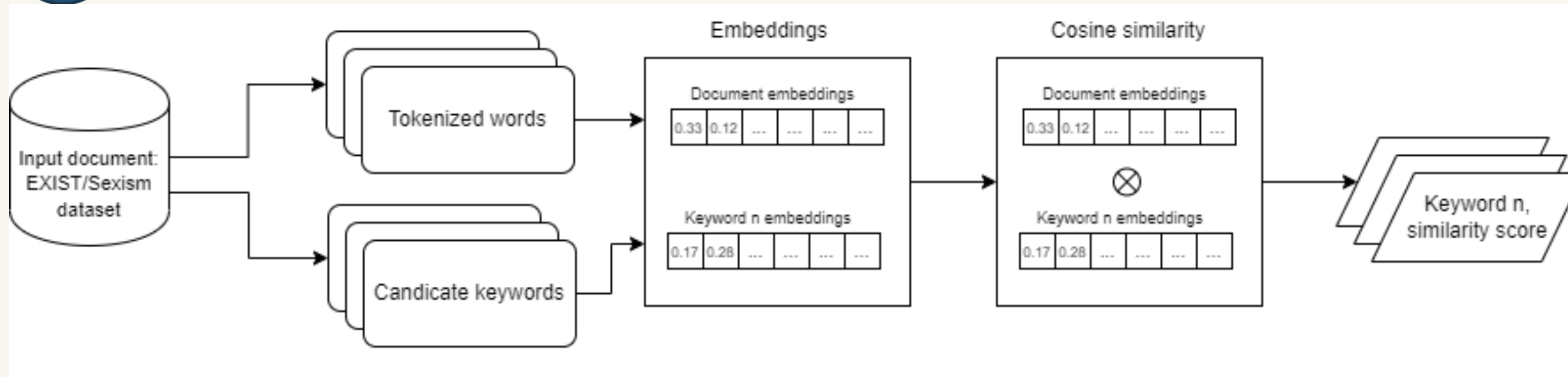
# 06 Results

## Manual annotation of data:

- 2 annotators
- 2 final categories:
  - Gender-related - posts that contain clear discussions about gender inequality, sexism, and other evident mentionings of these topics.
    - Potentially gender-related - posts that contain keywords although relatedness to the topic is uncertain (interim category).
  - Not gender-related - posts that while containing words related to gender are not about gender discrimination or sexism.
- Interrater reliability (Cohen's kappa coefficient) - 0.68

Manually review and annotate posts

Labelled dataset

119 Gender posts     3992 Non-gender posts
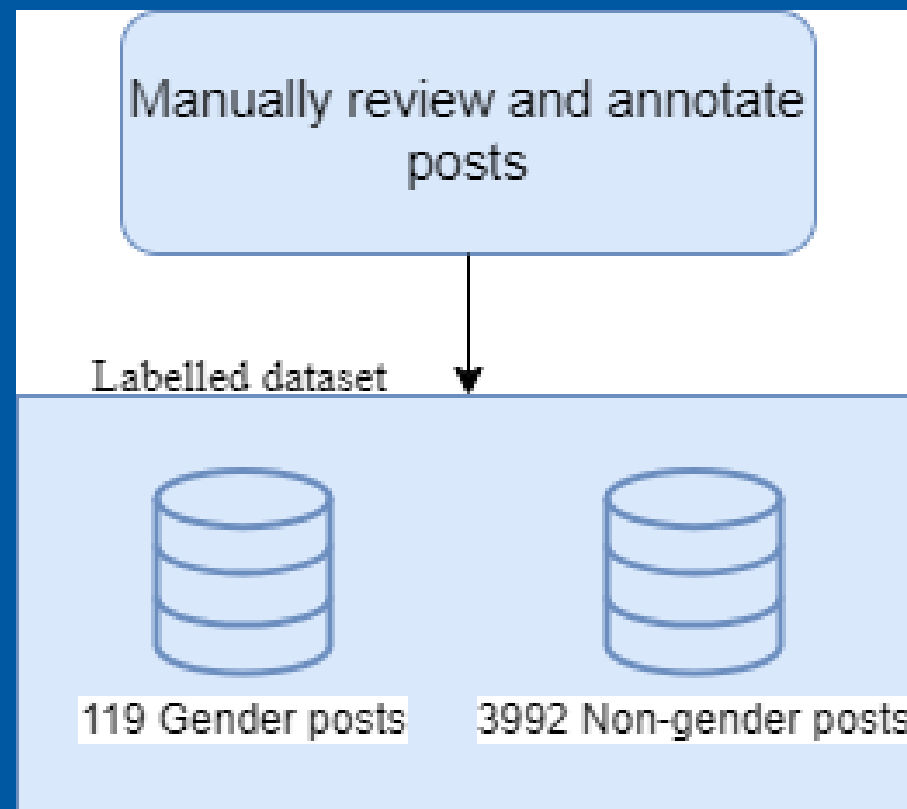
# 06 Results

Categories of gender discussions:

- Open coding method from the grounded theory was used
- 6 categories and the most common trends were identified:
    - Advertisement
    - AI
    - App Features
    - Company Policy and Censorship
    - Community
    - Content

| Raw text | Pre-processed text |
|----------|-------------------|
| Your filter is extremely sexiest and discriminatory,"A man in a swimsuit" is not a blocked prompt but ""a woman in a swimsuit"" is blocked"" a man having a shower"" not blocked ""a woman having a shower"" blocked.""A womans chest"" blocked ""a man's chest"" ....? I don't know I just got blocked for the next hour lol | 'filter', 'extremely', 'sexy', 'discriminatorya', 'man', 'swimsuit', 'block', 'prompt', 'woman', 'swimsuit', 'block', 'man', 'shower', 'block', 'woman', 'shower', 'blockeda', 'womans', 'chest', 'blocked', 'man', 'chest', 'dont', 'know', 'get', 'blocked', 'next', 'hour', 'lol' |
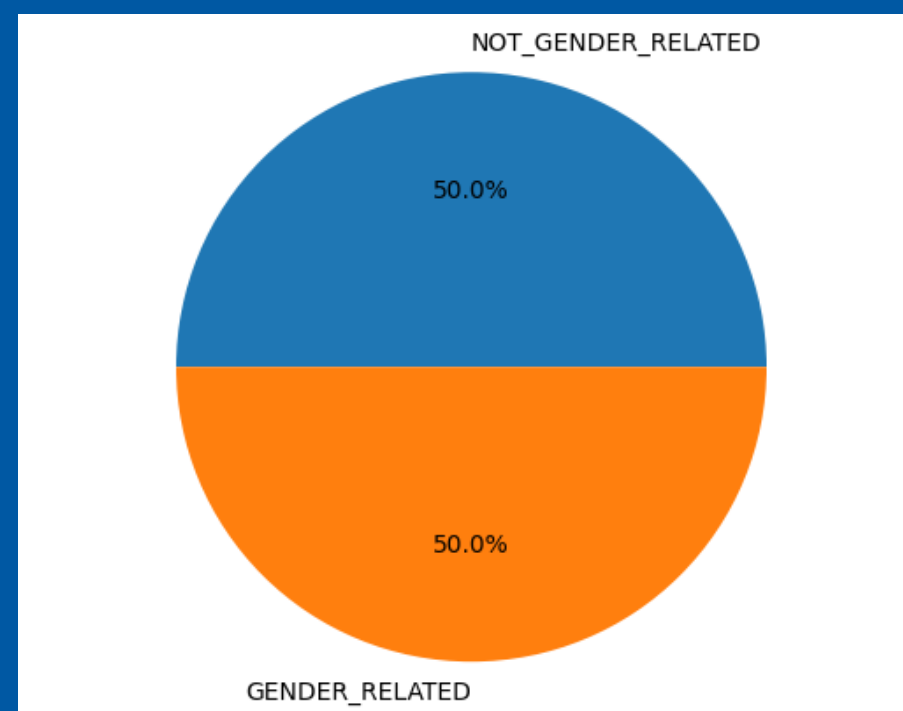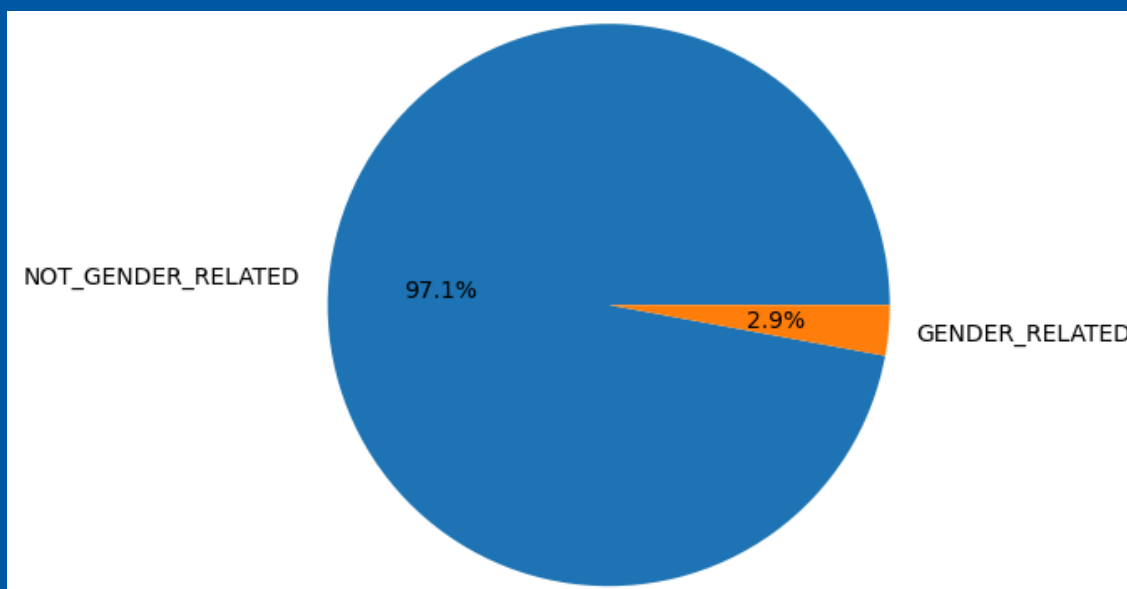




# 06 Results

## Automated classification:

- Data pre-processing
  - normalization, tokenization, and lemmatization
- Feature extraction
  - BoW, TF-IDF
- Addressing the class imbalance in Reddit data
  - SMOTE
- Classification models
  - SVM, LR, DT, NB, DistilRoBERTa, Llama-3-8B)
- Evaluation
  - Precision, Recall, F1-score, AUC

# 06 Results

ROC_AUC:

SVM

LR

DT

NB
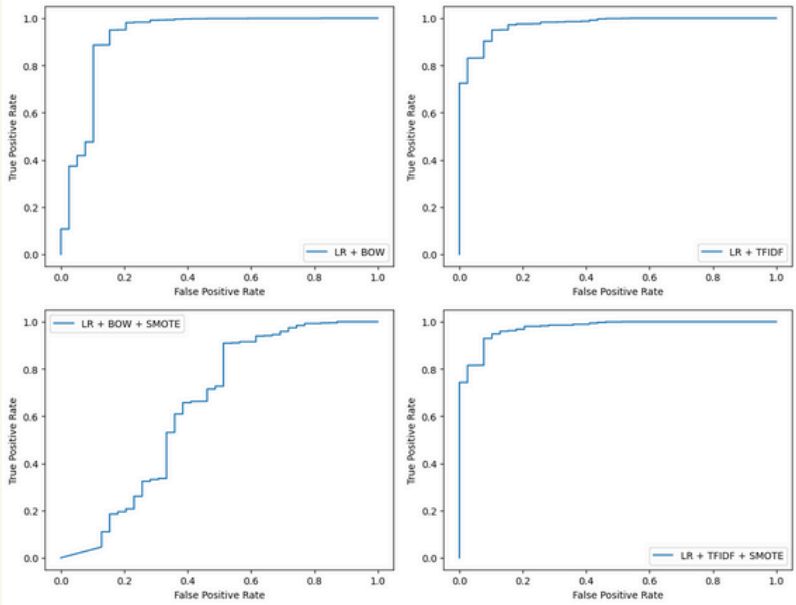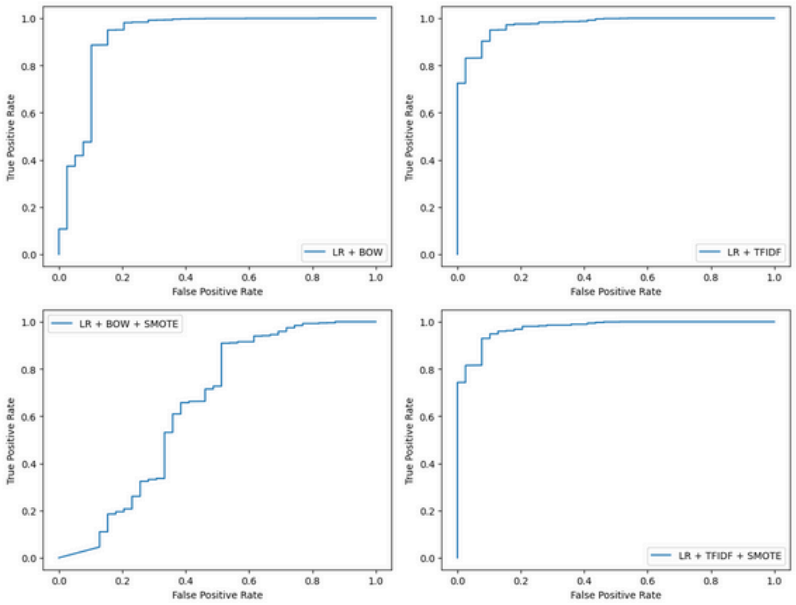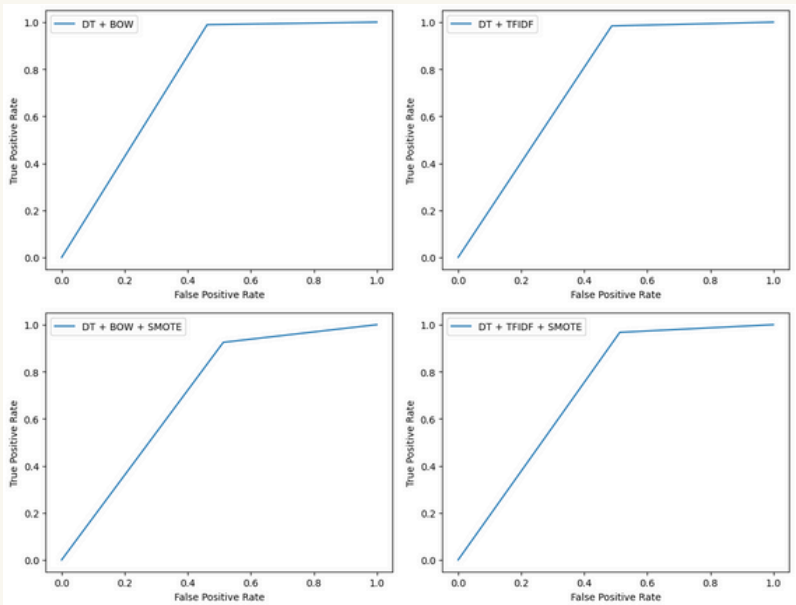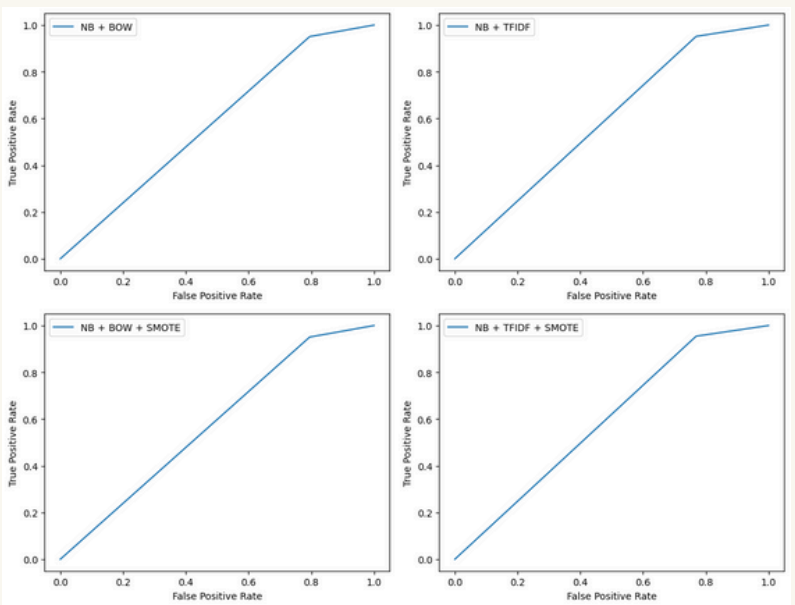
DistilRoBERTa
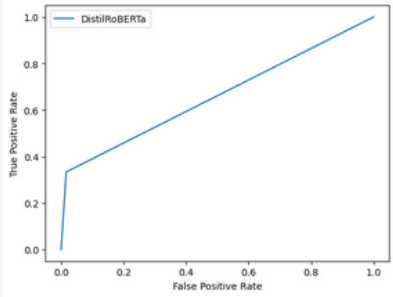
Llama-3-8B
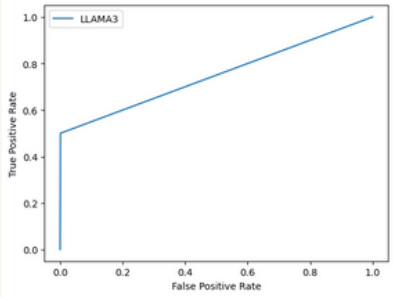
# 07 Conclusions

- Reddit data provides valuable information for investigating gender-related issues in mobile applications.
- Reddit data can be classified and categorized into separate groups of gender inclusivity violations in software products.
- Created a gender-related dataset for automated classification.

# 07 Conclusions

Answer to RQ1:

Yes, Reddit data provides meaningful information about gender-related discussions in mobile applications.
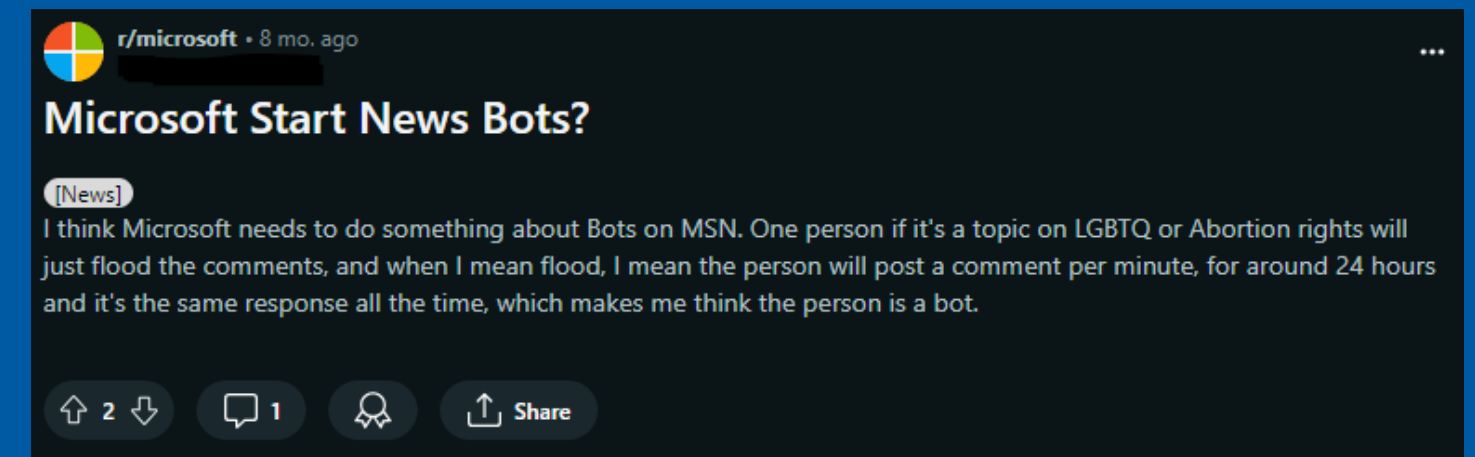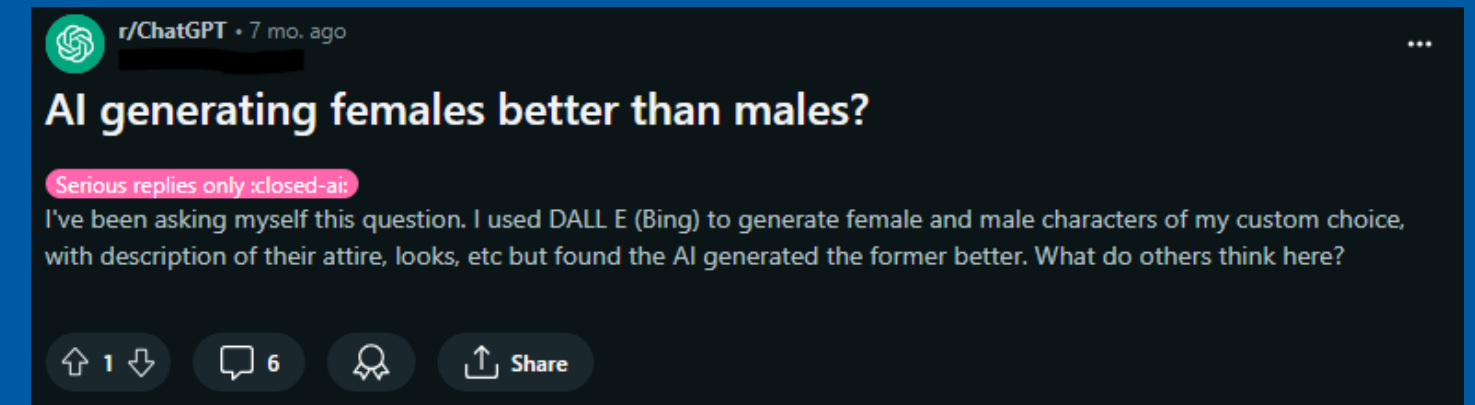
# 07 Conclusions

**Answer to RQ2:** Yes, we can further categorize the gender-related discussions in Reddit posts.

| Category | Topics |
|---|---|
| Advertisement | Gender representation and identity in digital platforms |
| | Inappropriate and sexually explicit advertising in online content |
| | Algorithmic bias and content curation on online platforms |
| AI | Gender bias and misinterpretation in language models |
| | Gender representation and identity in digital platforms |
| | LGBTQIA+ discrimination |
| | Navigating discussions on gender identity and online moderation |
| | Sexual objectification of women |
| App features | Account settings |
| | Video and conference filters |
| | Search algorithms |
| | Privacy and safety measures |
| | Emojis |
| Community | Hate speech and toxicity in online communities |
| Company Policy and Censorship | Gender-biased content moderation on social media platforms |
| | Censorship and content restrictions in online research platforms |
| | Accountability of online platforms in addressing sexual misconduct |
| Content | Algorithmic bias and content curation on online platforms |
| | Sexist, inappropriate or fraudulent content |

# 07 Conclusions

Answer to RQ3:

How do different ML models used for the automated classification of Reddit data for gender discussions in mobile applications compare with each other:

| Configuration | NOT GENDER RELATED | | | | | GENDER RELATED | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | P | R | F1 | ACC | AUC | P | R | F1 | ACC | AUC |
| BOW + SVM | 0.99 | 0.99 | 0.99 | 0.98 | 0.91 | 0.78 | 0.64 | 0.70 | 0.98 | 0.91 |
| BOW + LR | 0.99 | 1.00 | 0.99 | 0.99 | 0.92 | 0.88 | 0.56 | 0.69 | 0.99 | 0.92 |
| BOW + DT | 0.98 | 0.99 | 0.99 | 0.98 | 0.75 | 0.61 | 0.49 | 0.54 | 0.98 | 0.75 |
| BOW + NB | 0.98 | 0.95 | 0.96 | 0.93 | 0.57 | 0.11 | 0.21 | 0.14 | 0.93 | 0.57 |
| TFIDF + SVM | 0.98 | 1.00 | 0.99 | 0.98 | 0.95 | 1.00 | 0.33 | 0.50 | 0.98 | 0.95 |
| TFIDF + LR | 0.97 | 1.00 | 0.99 | 0.97 | 0.97 | 1.00 | 0.03 | 0.05 | 0.97 | 0.97 |
| TFIDF + DT | 0.99 | 0.98 | 0.98 | 0.97 | 0.76 | 0.48 | 0.54 | 0.51 | 0.97 | 0.76 |
| TFIDF + NB | 0.98 | 0.95 | 0.96 | 0.93 | 0.59 | 0.12 | 0.23 | 0.16 | 0.93 | 0.59 |
| BOW + SVM + SMOTE | 0.98 | 0.93 | 0.95 | 0.91 | 0.55 | 0.11 | 0.31 | 0.16 | 0.91 | 0.55 |
| BOW + LR + SMOTE | 0.98 | 0.95 | 0.96 | 0.93 | 0.64 | 0.16 | 0.31 | 0.21 | 0.93 | 0.64 |
| BOW + DT + SMOTE | 0.98 | 0.95 | 0.96 | 0.93 | 0.70 | 0.11 | 0.21 | 0.14 | 0.93 | 0.70 |
| BOW + NB + SMOTE | 0.98 | 0.95 | 0.96 | 0.93 | 0.57 | 0.11 | 0.21 | 0.14 | 0.93 | 0.57 |
| **TFIDF + SVM + SMOTE** | 0.99 | 1.00 | 0.99 | 0.99 | 0.95 | 0.91 | 0.54 | 0.68 | 0.99 | 0.95 |
| TFIDF + LR + SMOTE | 0.99 | 0.99 | 0.99 | 0.98 | 0.97 | 0.64 | 0.60 | 0.62 | 0.98 | 0.97 |
| TFIDF + DT + SMOTE | 0.99 | 0.96 | 0.97 | 0.95 | 0.76 | 0.30 | 0.54 | 0.39 | 0.95 | 0.76 |
| TFIDF + NB + SMOTE | 0.98 | 0.95 | 0.97 | 0.93 | 0.59 | 0.13 | 0.23 | 0.17 | 0.93 | 0.59 |
| DistilRoBERTa | 0.98 | 0.98 | 0.98 | 0.96 | 0.65 | 0.40 | 0.30 | 0.34 | 0.96 | 0.65 |
| Llama-3-8B | 0.99 | 1.00 | 0.99 | 0.99 | 0.75 | 0.75 | 0.43 | 0.55 | 0.99 | 0.75 |

# 07 Conclusions

## Limitations

Reddit API posts retrieval limit

Potential bias in manual annotation of data

Imbalanced training dataset

Small ratio of gender-related posts

## Future work

→ Increase the size of the dataset by collecting more data from Reddit or using paid 3rd party APIs

→ Involve more annotators to increase interrater reliability

→ Explore cost-effective machine learning models to overcome the class imbalance during training

Explore categories of gender-related discussions and the preparation of a multilabel gender dataset

# 08 References

[1]     Mojtaba Shahin, Mansooreh Zahedi, Hourieh Khalajzadeh, and Ali Rezaei Nasab. A study of gender discussions in mobile apps. arXiv (Cornell University), 1 2023.

[2]     Tahira Iqbal, Moniba Khan, Kuldar Taveter, and Norbert Seyff. Mining Reddit as a new source for software requirements. In 2021 IEEE 29th International Requirements Engineering Conference (RE), pages 128–138, 2021.

[3]     Humphrey O. Obie, Waqar Hussain, Xin Xia, John Grundy, Li Li, Burak Turhan, Jon Whittle, and Mojtaba Shahin. A First Look at Human Values-Violation in App Reviews. In IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS), pages 29–38, 2021.

[4]     Mattia Fazzini, Hourieh Khalajzadeh, Omar Haggag, Zhaoqing Li, Humphrey Obie, Chetan Arora, Waqar Hussain, and John Grundy. Characterizing human aspects in reviews of COVID-19 apps. In Proceedings of the 9th IEEE/ACM International Conference on Mobile Software Engineering and Systems, MOBILESoft '22, page 38–49, New York, NY, USA, 2022. Association for Computing Machinery.