Towards more Inclusive Software: A Large Scale Analysis of Inclusiveness from User Feedback

by

Nowshin Nawar Arony
B.Sc., Ahsanullah University of Science and Technology, 2018
A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Computer Science

Towards more Inclusive Software: A Large Scale Analysis of Inclusiveness from User Feedback

by

Nowshin Nawar Arony
B.Sc., Ahsanullah University of Science and Technology, 2018

Supervisory Committee

Dr. Daniela Damian, Supervisor
(Department of Computer Science)

Dr. Neil Ernst, Departmental Member
(Department of Computer Science)

ABSTRACT

In an era of rapidly evolving software usage, addressing the diverse needs of users from around the world has emerged as a critical challenge. Diverse users bring forth diverse requirements, encompassing factors such as human values, ethnicity, culture, educational background, technical expertise, preferences, personality traits, emotional states, and mental and physical considerations. Among the various aspects, inclusiveness, representing a core human value, is often unknowingly neglected during software development, leading to user dissatisfaction. Online platforms, such as forums and social media, offer users a space to express their opinions regarding a software. As a result, in recent times, software companies have recognized these platforms as a source of user feedback. Therefore, in this study, I leverage user feedback from three popular online sources: Reddit, Google Play Store, and Twitter (now known as X) to explore the inclusiveness related concerns from end users. I collected user feedback from the three sources for 50 of the most popular apps in the world. The 50 apps are selected from 5 types of software: business, entertainment, financial, e-commerce, and social media. I employed a Socio-Technical Grounded Theory approach and manually analyzed 23,107 posts across the three sources. Through this process, I identified 1,211 inclusiveness related posts. The research resulted in the development of a taxonomy for inclusiveness comprising 6 major categories: Fairness, Technology, Privacy, Demography, Usability, and Other Human Values. Along with that, I investigated the process of automatically identifying inclusiveness and non-inclusiveness related posts using 5 popular deep learning-based models. Upon experimenting with five deep learning models, I found that GPT-2 performed best on Reddit, achieving an F1-score of 0.838, BERT on the Google Play Store with an F1-score of 0.849, and BART on Twitter with an F1-score of 0.930. My research provides a detailed view of inclusiveness-related user feedback, enabling software practitioners to gain a more holistic understanding of such user concerns. The insights from this thesis can guide software organizations to increase awareness and address the inclusiveness aspects relevant to their product from an end-user perspective. I further provided implications and suggestions that can be used to bridge the gap between user values and software so that software can truly resonate with the varied and evolving needs of diverse users.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

Daniela Damian, for her invaluable guidance during my graduate studies. Her faith in my capabilities has been a constant source of inspiration to pursue my research endeavours.

Zane Li, for being the most incredible friend and mentor throughout my graduate journey. For always motivating me, especially during tough times. His unwavering support was instrumental in the success of this study.

Kezia Devathasan, for her extraordinary friendship, constant encouragement and making me feel at home.

Manish Sihag for his friendship and the wonderful moments we shared.

My parents and brother, for being the wind beneath my wings and continuously providing support and prayers for me.

Faiza, Farial and **Mohua**, for always being there and cheering for me.

My amazing colleagues from the **SEGAL Group**, for all the fun lunchtime conversations.

The INSPIRE program, for enabling me to learn valuable life lessons.

Those who do good in this world will have a good reward.
- Quran (Surah Az-Zumar, 39:10)

DEDICATION

To my parents for always inspiring and believing in me to pursue my dreams.

Chapter 1

Introduction

As software usage continues to grow worldwide, an increasingly diverse user base is engaging with these applications. The diverse group includes individuals from various genders, regions, ages, cultures, socio-economic backgrounds, political beliefs, people with physical and cognitive abilities, values, and educational backgrounds, among many others. However, software is often built for the "average user" [38] and fails to adhere to the diverse user needs. For instance, Twitter (currently known as X), a widely used social networking app with over 390 million global users [42], released an image cropping algorithm that automatically cropped images. It focused on important aspects of the picture, like faces and text, to optimize space on the main feed and allow multiple pictures in a single tweet. However, users soon identified that the algorithm could only detect white faces and cropped out faces of black people [11]. The topic soon became trending as thousands of users engaged in discourse about it. Historically, numerous similar incidents have emerged from user feedback on online platforms (i.e., Crowd), highlighting the lack of inclusiveness in software.

Inclusiveness is part of a broader category of human values, which relates to "what people hold important in their life" [40, 31]. In the context of software engineering, inclusiveness indicates the value that a user expects to be present in the software [31]. Moreover, inclusive software refers to "running anywhere" and "used by anyone" as coined by Savidis and Stephanidis [38]. Nevertheless, software companies often fail to make software inclusive during the initial phases of development. The functionalities of a software are typically determined during the requirement elicitation phase with the mindset of designing for an "average user". Software companies often follow a CI/CD (continuous integration and continuous delivery/continuous deployment) pipeline, which involves regularly eliciting requirements from users based on their

software usage. Users often leave feedback on online sources that companies use to identify further areas of improvement. Hence, these sources can be leveraged to identify and improve the inclusiveness aspect of software.

Crowd Requirement Engineering (CrowdRE) research has become a popular area of study for identifying product relevant information as large volumes of user feedback are received from various online platforms such as app stores, social media, and forums. Thus, CrowdRE methods can be leveraged to understand the aspects of inclusiveness. A growing body of research in "end user human aspects" has attempted to address and understand aspects like gender and accessibility using CrowdRE sources such as App reviews [41, 4]. Studies have further used Google Play Store to understand the human aspect relevant user feedback for 12 open source apps [20]. Therefore, in this research, I leverage the power of CrowdRE to investigate inclusiveness related feedback from the end users.

I employed a Socio-technical grounded theory (STGT) [18] approach to analyze user feedback collected for 50 of the most popular apps in the world from three sources: Reddit, Google Play Store, and Twitter. The 50 apps are selected from 5 distinct domains: business, entertainment, financial, e-commerce, and social media. The approach included manual analysis of 23,107 posts from all three sources, which resulted in the development of a taxonomy of inclusiveness related user feedback. I then used the labelled set to train deep learning models to automatically identify and generate inclusiveness related user feedback from the vast pool of feedback.

This thesis contributes to the growing body of research on inclusive software design and development. In the research, I uncover the inclusiveness related issues indicated by end users of popular for-profit apps. These apps are used by millions of users worldwide, providing a broad and diverse set of opinions regarding the apps. Therefore, through my research, I shed light on the user concerns that need to be addressed to make software more inclusive to users around the world.

1.1 Motivation

Inclusive software involves creating designs with everyone in mind and considering the full range of human diversity [2]. Traditionally, software requirements are gathered and incorporated into the software with a strong emphasis on technical aspects, frequently resulting in a neglect of diverse user needs. As a result, many users are unable to access certain features or, in some cases, are excluded from using the software

entirely. Such an issue may arise from a failure to address the problems related to inclusiveness. Thus, identifying and eliciting inclusiveness related issues encountered by users will help improve the user experience.

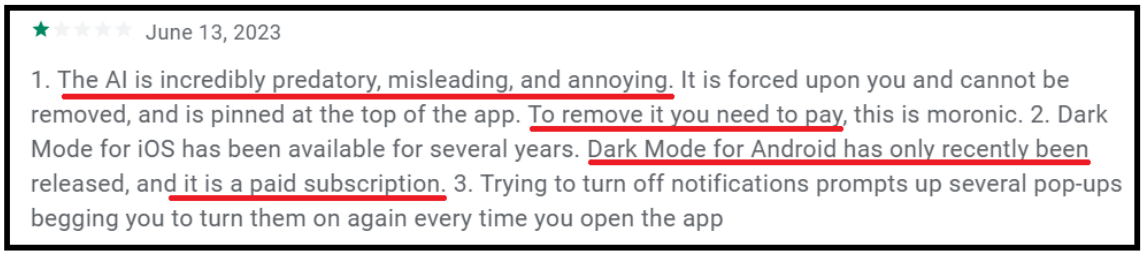

Figure 1.1: App review from Google Play Store. Underlined text indicates inclusiveness concern.

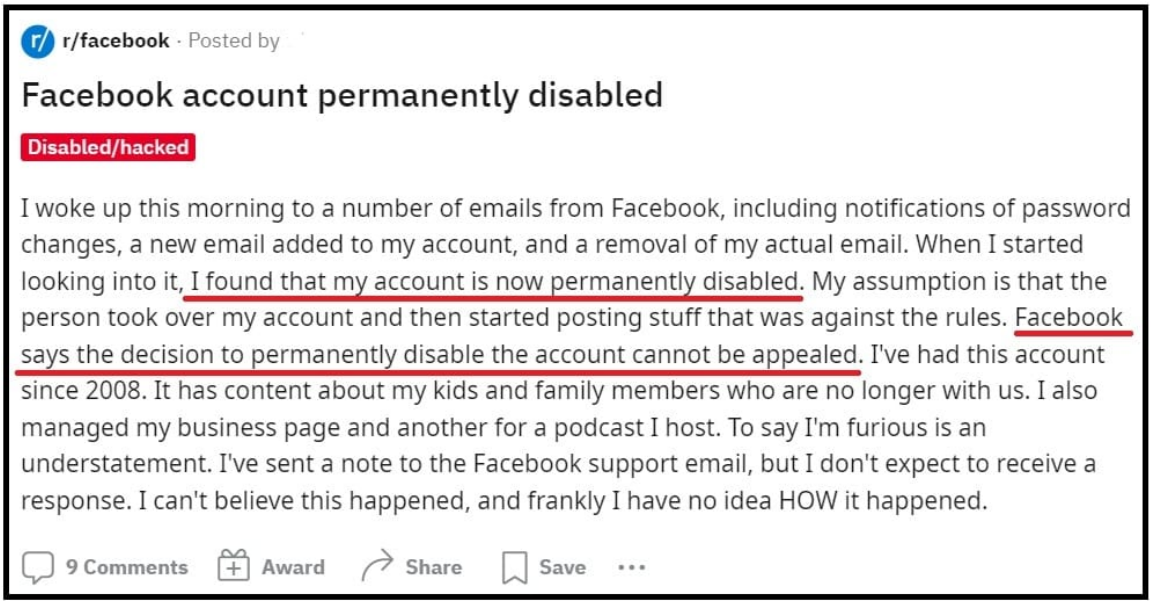

Figure 1.2: Reddit post from Facebook subreddit. Underlined text indicates inclusiveness concern.

User feedback from online platforms has become a popular source for gathering information about the software. App reviews, Reddit posts, or tweets offer users a space to express their opinions regarding the software. These sources serve as a rich source of information that can be useful in identifying problems that exclude users from the software. Figure 1.1 presents an example from Google Play Store where the user feels excluded from using dark mode on Android and removing the AI bot as they are not a paid subscriber. Similarly, the Reddit post in Figure 1.3 and tweet in Figure

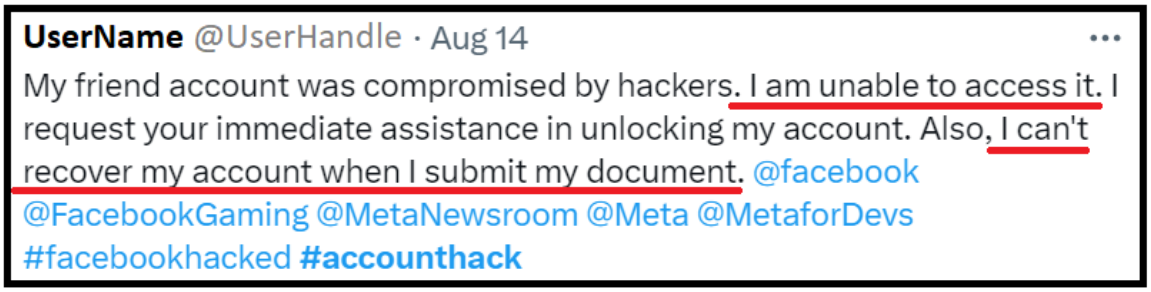

Figure 1.3: Tweet from Facebook user. Underlined text indicates inclusiveness concern.

1.2 demonstrates how Facebook disabled the user's account without any explanation and left no further options to appeal. The posts highlight an inclusiveness issue that Facebook and Snapchat should consider so users are allowed to use the platform without facing harsh consequences. In the examples, the users experience exclusion from the software as they are unable to access their accounts entirely. Overlooking such issues can eventually lead to the exclusion of a significant number of users from leveraging the full potential of the software. To further illustrate, consider an instance where an app does not support multiple languages. In such cases, it excludes users who do not understand the provided language. Many such issues can be identified from online user feedback that companies should take into account to ensure a more inclusive user experience.

Often, these issues stem from a lack of awareness, and the examples show that there is a need to address inclusiveness concerns. As illustrated in the examples, online sources carry insights into the various inclusiveness problems encountered by the diverse users of the software. Software companies frequently roll out features and changes that may cause such issues. Gaining more knowledge about the inclusiveness issues from the end user perspective can raise awareness and help design more inclusive software.

1.2 Problem Statement

The need to increase awareness regarding inclusiveness issues in software development is evident. Previous work examined user feedback for 12 open source apps, revealing only 31 relevant posts on inclusiveness, and categorized them into five categories [20]. Furthermore, open source apps often lack large user bases and profitability (i.e.,

underlying usage numbers), which is likely less important than for-profit apps. Therefore, there remains ample room for exploration of the inclusiveness category from a more diverse and larger user base perspective. A better understanding of inclusiveness issues is critical to raise awareness and facilitate the design and development of more inclusive software solutions.

1.3 Research Question

With the goal to shed more light on the inclusiveness issue, I embark on this research that is guided by the following research questions:

RQ1 What are the different types of inclusiveness related user feedback found on online sources?

RQ2 How does inclusiveness related user feedback differ for different types of apps and sources of feedback?

RQ3 Can we automatically identify inclusiveness related user feedback from online sources?

1.4 Methodology

My research aims to gain a deeper understanding of inclusiveness related user feedback from end users and report a classifier to automatically identify such feedback. To achieve this goal, I first use a Socio-technical grounded theory (STGT) approach [18] to conduct a qualitative analysis of the user feedback. To conduct the analysis, I analyzed user feedback obtained from 50 of the top apps in the world, which contain millions of users.

As sources, I choose app reviews from Google Play Store, posts from Reddit, and tweets from Twitter, considering the wide popularity of the sources across the world. The three platforms have previously been studied and indicated as useful in providing app relevant information such as bugs and features [47, 26, 17]. Previous research studies have also shown that app reviews can provide information on human aspect relevant concerns [20]. In addition, I experiment with five deep learning models to automatically identify the inclusiveness related user feedback and report an auto-

mated classifier that yields the best results for each source. I describe each step of the methodology in more detail in Chapter 3.

1.5 Research Contributions

My research provides the following contributions:

- Empirical evidence of the most prevalent inclusiveness categories in the form of a Taxonomy based on user feedback from top 50 popular for-profit apps containing millions of users from three major sources: Reddit, Play Store, and Twitter.
- Empirical insights into which category is more prevalent in which type of app and source of feedback.
- A manual annotated dataset of inclusiveness user feedback that future researchers and practitioners can use.
- An automated classifier to identify inclusiveness related user feedback from the three sources that companies and practitioners can leverage to address the inclusiveness concerns.

1.6 Thesis Outline

I organized the thesis as follows:

Chapter 1: Introduction

This chapter begins with introducing the research topic, presents the motivation and problem statement for the research, includes an overview of the research methodology, states the research questions, and unveils the contributions of the research.

Chapter 2: Background and Related Work

This chapter explores the relevant background information and previous studies related to the research topic.

Chapter 3: Methodology

This chapter describes the research methodology in detail, including the data collection process, the selection of variables, the research design, and the analytical techniques used.

Chapter 4: Research Findings

This chapter presents the findings of the research, showcasing the results obtained from the data analysis. It highlights the key discoveries, patterns, and trends identified during the research process and provides an overview of the research outcomes.

Chapter 5: Discussion and Implications

This chapter discusses the research findings in depth, considering their implications and significance. It explores the implications of the research results for theory, practice, and future research directions.

Chapter 6: Conclusion

This chapter includes the final remarks about the research.

Chapter 2

Background and Related Work

To better understand the aspects of inclusiveness, I conducted an in-depth literature review. In this section, I describe the existing literature on inclusive software and user feedback, along with the research gap in relation to the literature.

2.1 Inclusive Software

The term inclusive software relates to the notion of "universal access" which implies software that is accessible and usable by everyone [38]. The underlying philosophy behind designing an inclusive product is to ensure that the product can be used by as many different users as possible rather than excluding anyone [35]. Pattison and Stedmon [35] suggest that when a product development process incorporates a human-centric approach with a focus on effectiveness and adaptability alongside an inclusive design approach, it helps prioritize inclusion for a diverse user base. A study by Savidis and Stephanidis [38] highlights the significance of providing the necessary tools to support inclusive software design and development. The authors indicate that an important aspect of inclusive software development is identifying user requirements that emerge from interaction with the software.

Conventionally, software is developed with a focus on the average user, and requirements for the software are developed with this perspective [38]. Although software is primarily intended to be neutral, software interfaces often contain stereotypical visual components that negatively impact many users' sense of belonging [28]. Metaxa-Kakavouli *et al.* [28] conducted a study where they introduced students to two webpages with identical content but different interfaces, one with more mascu-

line features, while the other was gender-neutral. The results indicated that women displayed less interest in using the masculine interface page and engaging with its content. In another study, Burnett *et al.* [6] revealed that problem-solving software is developed with the perception that users will adopt the features through tinkering. However, these features are statistically preferred by men over women, making the software less inclusive for women. These studies highlight the need for more inclusive software.

To make the software more inclusive, Nunes *et al.* [29] proposed a conceptual model for gender-inclusive requirements that involves creating a gender-inclusive requirements document. The document supports practitioners in integrating the model into the requirement elicitation process. Upon evaluation of the model, they found 83.9% positive response in terms of the usefulness of the model [30]. The GenderMag (Gender Inclusiveness Magnifier) method developed by Burnett *et al.* [6] uses personas encapsulating five facets of gender differences to analyze gender inclusivity in software. An empirical investigation of GenderMag identified biases in industrial software and helped derive design changes that improved the inclusiveness of the software [45]. Guizani *et al.* [16], in their study, introduced a Why/Where/Fix methodology for identifying and rectifying inclusivity bugs within an Open Source Software project. They observed a significant 90% reduction in inclusivity bugs due to implementing their approach.

While the studies focused on addressing gender-related concerns, the concept of inclusion extends beyond gender. In an extension of the GenderMag study, Mendez *et al.* [27] introduced InclusiveMag, a three-step approach involving Scope, Derive, and Apply, aimed at evaluating various aspects of diversity, including cognitive abilities, vision, educational background, socio-economic status, and age. The authors concluded that facilitating this type of assessment of software inclusivity in the early stages of development is critical in addressing the underlying biases present in software. Moreover, recent studies have also indicated that to make software more inclusive, software companies need to better understand human aspects such as age, emotions, personality, human values, gender, ethnicity, and culture [15, 14]. There are various ways to understand the different human aspects of diverse users. For example, co-design or participatory design techniques where users are invited to participate and provide feedback during the design process [9]. However, as software grows and becomes more prevalent around the world, it becomes difficult for companies to conduct such design sessions to understand their diverse users' needs. In such

cases, CrowdRE (Crowd Requirement Engineering) [12] techniques, such as analyzing user feedback from online sources can be leveraged.

2.2 User Feedback

Prior literature has shown that user feedback is beneficial for the continuous improvement of software quality [33]. User feedback from the online platforms, i.e., "Crowd" offers a large set of diverse end users [13], that has been studied to identify a variety of user needs. Requirement-related information such as feature requests, bugs, and user experiences is one the popular forms of insights that are identified from app reviews [26], Twitter (now referred to as X) [47], and Reddit [19].

More recently, Fazzi *et al.* [10] analyzed 2,611 app reviews from 57 COVID-19 apps and found nine categories of human aspect related discussions that impact software usage. The authors implied that these human aspects are not always taken into consideration and should be addressed during development. Another research on android apps more focused on accessibility issues revealed that the majority of these apps contain significant problems that prevent individuals with disabilities from using the apps [4]. The study demonstrates that various subcategories of human aspects are identifiable from user feedback, which can raise awareness amongst developers and companies, enabling them to incorporate these insights during development. Similarly, Shahin *et al.* [41] conducted an analysis of gender related discussions on app reviews and found six major categories: AppFeatures, Appearance, Content, Company Policy and Censorship, Advertisement, and Community. In addition, they automated the identification of gender and non-gender related discussions and acquired an F1-score of 90.77%.

Li *et al.* [24] obtained 4.5 million posts from Reddit and found 9 significant topics related to privacy concerns. Likewise, Olson *et al.* [32] examined 586 subreddit communities and identified discussions on ethical concerns from end users regarding social platforms. These studies provide empirical evidence that Reddit is a valuable source for gathering and understanding user concerns.

2.3 Research Gap

Previous research by Khalajzadeh *et al.* [20] characterised human-centric issues into three major categories: App Usage, Inclusiveness, and User Reaction through manu-

ally examining 1,200 app reviews and 1,200 GitHub issue comments for 12 open source projects. Their work highlights how human-centric issues are addressed within the same projects by both end-users and developers. The authors present a taxonomy for human-centric issues and employ machine learning and deep learning models to automatically classify the discussions. Figure 2.1 shows the inclusiveness categories identified in their work.

However, the inclusiveness category reported in their study is representative of only open source app users, which lacks a more diverse range of users. The study further uses app reviews as a source and presents only 31 inclusiveness related posts, which is insufficient to represent the diverse end user concerns related to inclusiveness. Hence, to address the gap, I conduct a deeper analysis of inclusiveness from an end-user perspective, leveraging three popular sources of user feedback. A thorough analysis of user feedback from three different sources for popular for-profit apps will enable me to triangulate and identify the inclusiveness related needs of a diverse users.

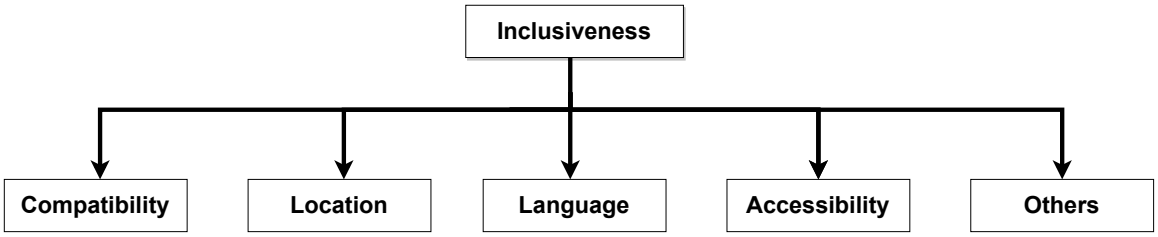

Figure 2.1: A taxonomy of inclusiveness category proposed in previous work by Khalajzadeh *et al.* [20]

Chapter 3

Methodology

To address the need to gain a deeper understanding of inclusiveness through the lens of end-users, I conducted an exploratory research. The thesis was guided by the following research questions:

RQ1 What are the different types of inclusiveness related user feedback found on online sources?

RQ2 How does inclusiveness related user feedback differ for different types of apps and sources of feedback?

RQ3 Can we automatically identify inclusiveness related user feedback from online sources?

I began by collecting user feedback posts from three popular platforms: Reddit, Google Play Store, and Twitter (presently known as X). To analyze the data, I used the Socio-Technical Grounded Theory (STGT) [18] approach. In particular, I applied the STGT for data analysis method that includes open coding, constant comparison, and basic memoing. I then leveraged five state of the art deep learning models to identify and propose a method to automatically classify inclusiveness related user feedback. My research aims to provide empirical evidence on the different categories of inclusiveness related user feedback and an automated classifier to identify the feedback. Figure 3.1 presents the overview of the methodology adopted in this thesis. The methodology includes data collection, random sampling, and applying STGT techniques for data analysis and labelling. After acquiring the labelled dataset, I presented the findings in a structured manner and use the labelled set to build automated classifiers.

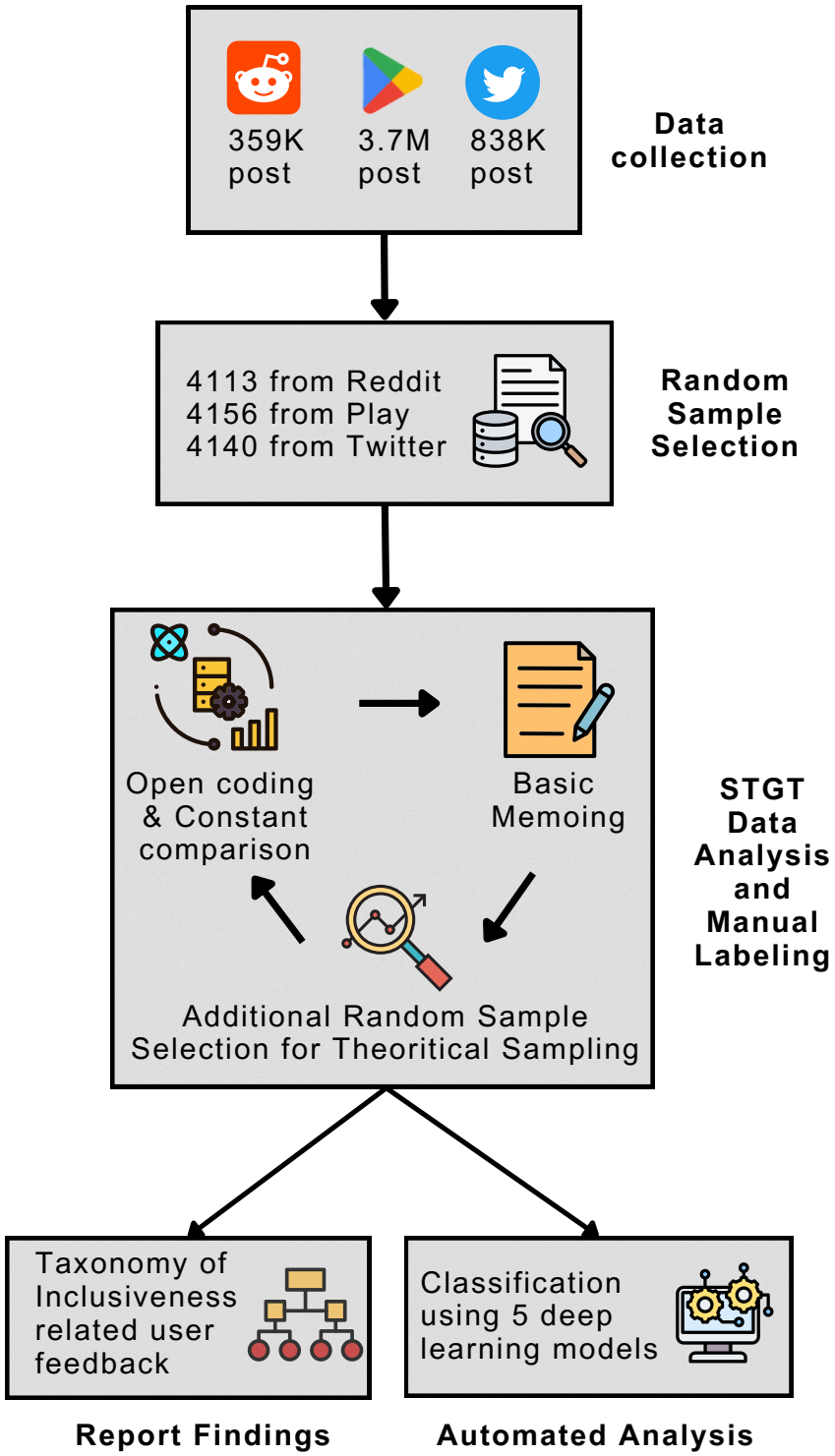

Figure 3.1: Overview of the research methodology

3.1 Data Collection

To achieve a broad perspective of the end user feedback, I collected data from three popular online sources of user feedback: Reddit, Google Play Store, and Twitter. These choices are based on prior studies that have successfully uncovered software related insights from these channels, such as bugs and features [26, 47, 19]. Reddit stands out for its distinctive feature of allowing users a high character limit, enabling them to engage in elaborate discussions. Within a single Reddit post, users have the space to provide elaborate details about a particular issue, a level of depth that is often unavailable in other comparable feedback sources. The Google Play Store, on the other hand, offers app users the opportunity to leave reviews about any app they have used. This platform allows software organizations to elicit concerns and feedback about specific apps. It provides a direct line of communication between users and developers, facilitating the identification of issues and potential improvements. In contrast, Twitter, one of the most popular social media platforms, supports short-form textual user discourse on a wide range of topics. Despite its character limit, Twitter has proven to be a resourceful source of requirement-relevant information for organizations to analyze [17]. Users often share their thoughts, opinions, and feedback in concise yet impactful ways on Twitter. This multi-source approach ensures that I can capture a wide spectrum of insights about inclusiveness from the users.

To compile the data, I first developed a list of the 50 most popular apps from Google Play Store. I selected the top 10 apps from 5 different domains: business, entertainment, financial, e-commerce, and social media. Table 3.1 provides a list of the apps under each domain. Further details for each app (google play id, subreddit, and hashtag) are provided in Appendix A. To ensure I acquire the most up-to-date and relevant data, I specifically targeted apps that have received reviews between January 1, 2022, and December 31, 2022. These apps offer a diverse group of users from across the world. The diversity ensures that the data encompasses the perspectives of users with varying preferences, needs, and expectations. The same list is used to scrape the data for Reddit and Twitter as well, thereby giving us a unified range of apps.

To collect data from Reddit, I use a publicly available dataset [46] and obtain over 380,000 reddit posts. Next, I collected 9,000,000 app reviews from Google Play using the library Google Play Scraper [1]. Lastly, I use the snscrape library [2] to scrape

[1] https://github.com/JoMingyu/google-play-scraper
[2] https://github.com/JustAnotherArchivist/snscrape

838681 discussions from Twitter.

After collecting all the data, I proceeded to filter the original data by first removing any empty posts. Additionally, I filter out any post that has less than three words as I believe that posts that cannot satisfy this criterion most likely do not provide meaningful information. After the filtration, the dataset comprised over 3.7 million app reviews, 838 thousand tweets, and 359 thousand Reddit posts.

Table 3.1: List of apps under each type

Business	Entertainment	Financial	E-commerce	Social Media
Adobe Acrobat	Amazon Prime Video	Crypto.com	Afterpay	BeReal
Fiverr	Disney+	Binance	Amazon Shopping	Discord
Lightroom	Hulu	Google Pay	Ebay	Facebook
LinkedIn	Netflix	Paypal	Temu	Messenger
Microsoft Outlook	Paramount+	CoinBase	Etsy	Instagram
Microsoft Teams	Peacock TV	Robinhood	Shopee	Snapchat
Trello	Pluto TV	Cash App	Shopify	Twitter
Webex	Roku	Venmo	Walmart	WhatsApp
Workday	Tubi	Wealthsimple	SHEIN	TikTok
Zoom	YouTube	Webull	Craigslist	Telegram

3.2 Qualitative Analysis

After collecting and preprocessing the user feedback, the next step of the methodology is analysis. I leveraged the Socio-Technical Grounded theory (STGT) [18] to analyze the data. STGT is a modern Grounded theory approach well suited for research in Software Engineering (SE). The author of STGT writes *"SE is neither exclusively social nor purely technical, rather, predominantly socio-technical (ST)"* [18]. STGT is

a modern Grounded theory approach well suited for research in Software Engineering (SE). It enables a more focused study through a lean (i.e., lightweight) literature review and the use of new data collection methods, such as data mining techniques from online sources. STGT can be applied in either its full form, which produces novel theories, or in a more limited manner, utilizing the basic data analysis techniques to establish important categories or initial hypotheses. Due to the exploratory nature of my thesis, I opted for the second option to analyze online user feedback. After all, my goal is to develop an understanding of inclusiveness related concerns from the perspective of end users.

The STGT for data analysis steps consists of open coding, constant comparison, basic memoing, and theoretical sampling. For the basic data collection step, I first collected data as described in Section 3.1. Then, I employed a random sampling technique with a 99% confidence level and a 2% margin of error. The random sampling allowed me to obtain a significant amount of data selected randomly from the user feedback for all 50 apps within each of the three sources. Using the technique, I acquired 4113 samples from Reddit, 4156 from Google Play Store, and 4140 from Twitter. This sampling approach guarantees samples across all 50 apps.

To guide the coding technique, in line with the STGT, I conducted a lightweight literature review in preparation for the study to identify the existing understanding of inclusiveness in software engineering (as outlined in the Background and Related Work section). I borrow the description of inclusiveness from Khalajzadeh *et.al* [20], and define inclusiveness related user feedback as *"any user feedback related to the inclusion, exclusion or discrimination toward an individual or a specific group of users while using the software."* In line with this definition, two members of the research team embarked on the task of reviewing the randomly selected data and assigning an inclusiveness or non-inclusiveness label to each post. Simultaneously, if a post is given an inclusiveness label, the team conducted an iterative, open coding strategy to identify the characteristics of each inclusiveness related user feedback. When a post was given an inclusiveness label, the annotators further included a code based on the characteristics of the feedback. Table 3.2 illustrates examples of the labelling process. For example, the quote *"Worse customer service. Don't any respect for people. Several times Amazon canceled my order without reason. I tried to wake them up about costumer rights, but they prefer to ignore rather than understand the concept. Amazon thinks there is no any consequences for them because this corporation already grows up and don't need to care about us. I really disappointed and sure they will fall*

in the end" was first labelled as *inclusiveness*-related. Additionally, the code *service* was assigned to it as the user expressed frustration about feeling excluded as they do not receive customer support.

Table 3.2: Example open coding of raw quotes and relationship to the taxonomy of inclusiveness

Raw Quote indicating an *inclusiveness*-related concern	Code (sub category)	Category
Worse customer service. Don't any respect for people. Several times Amazon canceled my order without reason. I tried to wake them up about costumer rights, but they prefer to ignore rather than understand the concept. Amazon thinks there is no any consequences for them because this corporation already grows up and don't need to care about us. I really disappointed and sure they will fall in the end.	Service	Fairness
Problem with subscribing to discord nitro, Help? I used to be a nitro subscriber and I had to upload my billing address, it wont recognize my location even though I've tried everything, I wish it would say something like address invalid or incorrect postal code would help a lot	Location	Demography
This app is horrible. You can't do or say anything that you wanna do or say. If you want to be censored and controlled, then this is the app for you. They want everybody to think the same & what's the fun in that? There is no diversity & no respect, just overbearing control & censorship.	Freedom	Other human values

Any time a new code emerges, the two annotators would meet and discuss the implications of the new. Constant comparison method is performed to compare the derived codes across all three user feedback sources to uncover the underlying patterns and relation between the codes. The team further used a memoing technique to document the reflections on the emerging codes and categories. The memos are presented in the Appendix C of this thesis. Memoing is a significant part of STGT data analysis, encapsulating the researchers' thoughts and enabling a systematic development

of categories from the initial codes.

Another important aspect of STGT is theoretical saturation, which means "when data collection does not generate new or significantly add to existing concepts, categories or insights, the study has reached theoretical saturation" [18]. Therefore, the two human annotators continued the labelling process until no new categories emerged and code definitions became stable (i.e. saturation). During this process, the initial random sample became insufficient to reach saturation, and I further randomly sampled additional posts from all three sources and analyzed them until reaching saturation for each category. The team stopped labelling when no additional insights emerged from the last 200 posts coded from Reddit and Google Play Store. For Twitter, due to the presence of irrelevant discussions in the data, saturation was reached after coding 500 additional posts.

In total, the team labelled 4,647 discussions from Reddit, 4,949 from Google Play Store, and 13,511 from Twitter. At this point in the labelling, I concluded that theoretical saturation is achieved. Following the above described process, I found categories of inclusiveness, which I present in Section 4.1 in the form of a two-layer *taxonomy of inclusiveness*. The taxonomy contains six categories forming the primary layer and the 18 codes distributed within each category as a sub-category. The labelled set is further used to examine a classifier to automatically identify inclusiveness related user feedback.

3.3 Automated Classification

Along with the establishment of a taxonomy of inclusiveness, the next step of my research included devising an automated technique to identify inclusiveness related concerns from a large amount of user feedback without intense human labour. The aim was to identify an automatic approach for identifying inclusiveness related feedback using classification.

I experiment with deep learning models that can automatically classify inclusiveness and non-inclusiveness user feedback. Recall in Section 3.2 that the research team collated a large human annotated set of user feedback. This labelled set serves as a ground truth to train a classifier.

To this end, I experimented with five widely recognized transformer based deep learning models that have demonstrated effectiveness in text classification. I used five state of the art deep learning models that are pretrained and fine tune the models

using the ground truth data and measure the performance of the classifiers. To train the model, I used a balanced set prepared with an equal number of inclusiveness and non-inclusiveness posts. The deep learning models I use in this research are:

- BART (Bidirectional and Auto-Regressive Transformers) [23]
- BERT (Bidirectional Encoder Representations from Transformers) [8]
- GPT-2 (Generative Pre-trained Transformer 2) [36]
- RoBERTa (Robustly Optimized BERT Approach) [25]
- XLM-RoBERTa (Cross-lingual Language Model - Robustly Optimized BERT Approach) [8]

For the evaluation, I used a set of test samples to measure the performance of the classifiers and used precision, recall, F1-score, and accuracy as the metrics. These metrics are used to evaluate the results for all 5 classifiers, enabling me to compare the performance of each model.

Chapter 4

Findings

The analysis of the data allowed me to uncover the underlying inclusiveness related concerns from the diverse users of the apps. The approach resulted in a large labelled dataset which enabled me to develop a taxonomy based on the recurring patterns in the dataset. I further analyzed the prevalence of the taxonomy categories in each type of app. Finally, I experimented with the 5 classifiers with the aim of developing an automated approach to classify the inclusiveness and non-inclusiveness related user feedback from large amounts of data. In this section, I present the findings of the research in light of the research questions.

RQ1 What are the different types of inclusiveness related user feedback found on online sources?

RQ2 How does inclusiveness related user feedback differ for different types of apps and sources of feedback?

RQ3 Can we automatically identify inclusiveness related user feedback from online sources?

4.1 Taxonomy of Inclusiveness

This section presents the answer to the first research questions: *What are the different types of inclusiveness related user feedback found on online sources?*

To investigate the answer to this questions, the research team manually annotated 4,647 posts from Reddit, 4,949 from Google Play Store, and 13,511 from Twitter using the STGT analysis into two classes: inclusiveness and non-inclusiveness, following the

definition provided in Section 3.2. At the same time, each of the inclusiveness related user feedback was labelled into sub-categories using the Socio-Technical Grounded Theory approach described in 3.2. To this end, the research team found 712 inclusiveness posts from Reddit, 377 from Google Play Store, and 116 from Twitter. The reason behind the smaller number of inclusiveness posts from Twitter is due to the presence of excessive noise in the data. The majority of posts are either ads or pertain to some social or political topic rather than the app itself. Nevertheless, I derived 6 categories of inclusiveness from the three sources. In addition, I also identifies sub-categories of inclusiveness that pertain to each of the 6 categories. I presented the categories and their sub-categories as a taxonomy for inclusiveness from end user feedback in Figure 4.1. I describe each category with examples identified from the 1,211 inclusiveness user feedback and include the percentage of the prevalence of each category from each source, where "R" indicates Reddit, "A" indicates app reviews from Google play store, and "X" symbolizes results from Twitter. Within each category, I include examples and explanations of the sub-categories.

4.1.1 Fairness

Fairness inclusiveness related user feedback are the most common category and account for almost a third of all inclusive discussions (R: 29.5%, A: 33.42%, X: 15.6%). The fairness category covers any user feedback where a user describes unfair behavior or treatment during app usage. Specifically, fairness includes issues such as encountering a ban or restricted access to features. I found that fairness usually concerns three sub-categories: ***Terms/Conditions***, ***Recommendation***, and ***Services***.

Terms/Conditions: I observed users frequently complaining about unjustified banning or restriction of their accounts.

(Reddit) - *"I got banned on YouTube because of malware on my [laptop] and they uploaded scam videos on my account and I got terminated. Ive got rid of the malware and I requested an appeal now they wont give me my account back."* ***(Youtube)***

(Twitter) - *"... I am limited still unable to follow. Shadowed or limited since 2015. #twitter"* ***(Twitter)***

(Play Store) - *"I didn't get a account waring they just permanently banned me without giving me a warning and it's not my fault kids kept saying "You're 7 8 10"*

and they don't get banned wow and I want my account back I lost 1k followers and it's not my fault so 1 star for now I try to download my data and make a new account that won't work" ***(TikTok)***

These quotes from the three sources exemplify similar types of issues: the perception that the software apps are not *fairly* treating end users and that there is little transparency for what transpired. In the YouTube example, the user feels it is unjustified to terminate their account for a malware they did not cause. In the Twitter example, the user describes a limited/shadow banned user experience, and in the TikTok example, the user describes an account ban without any kind of warning. The commonality between the behaviour of the apps is that account restrictions do not seem impartial. Furthermore, it seems in these scenarios, the apps also suffer from a lack of additional support for users to appeal their cases.

(Reddit) - *"Facebook disabled both me and my wife's accounts for infringing someone else's intellectual property rights. We never done anything wrong. We admin and manage several groups on Facebook in which members post goods etc. When Facebook send any notifications we go and check and if necessary we remove and ban those members on those specific groups."* ***(Facebook)***

(Play Store) - *"Amazing job sorting what's important and what's not, main inbox full of ads whereas spam folder full of work related documents, your algorithm nearly got me fired."* ***(Microsoft Outlook)***

(Play Store) - *"Venmo somehow allows a user to accidentally charge me a fee because they marked me as a seller and not just a user"* ***(Venmo)***

(Reddit) - *"So i reported the bots that mass spammed people and the fake link posing as the real account, and what does facebook do? Nothing, refused my report and said it didn't go against community guidelines. It's really weird a huge site like Facebook let scams go untouched, and with the system in place you cannot actually ask them directly or explain the situation. "* ***(Facebook)***

In these examples, I found further instances of users complaining about the perceived lack of fairness in their user experience. Account restrictions and limitations often stem from company policies and are carried out by automated decision-making algorithms, frequently in the form of bots. The feedback about Facebook are clear

examples of this, where accounts are marked as infringing on community guidelines. Software organizations often employ these bots to automatically resolve potential account issues and flag infringing accounts. However, it appears from the perspective of users, many feel that there exists algorithmic bias.

Recommendation I uncovered several users expressing frustration with the automated recommendations in the apps. I found this occurring frequently as software apps use machine learning to recommend services or content to users. Users start complaining when they feel that these machine learning systems fail to consider their preferences, which occurs when an app unexpectedly and continuously suggests content that the user does not want to see. When this happens and is not rectified by the app developer, it may cause the user to feel that the app is unfairly treating the user. The recommendations may pertain to content that users did not subscribe to or have marked as unwanted. In addition, they may be subjected to unwanted ads, leading to a feeling of exclusion and causing them to leave the app entirely.

💬 **(Play Store)** - *"Worst recommendations on the app. I am spending so much time to see videos on the Fb. I have noticed facebook recommends me the videos that i don't want to see... Every time i am blocking the video page but still it recommends."* ***(Facebook)***

💬 **(Reddit)** - *"... recently my [Timeline] is full of random reels and post from people I dont follow. Does anyone know how fix this in settings? Its very annoying and usually leads me to closing the app after about 5 minutes."* ***(Instagram)***

💬 **(Play Store)** - *"... this app forces you to see [specific news channel] content. You partner with these trash news sites but not enough good ones to round it out. Please stop pushing these machine generated news sites so hard on those who actively don't want it."* ***(Robinhood)***

💬 **(Twitter)** - *"... Im about to delete all other platforms because of the amount of adds. Also, dont really care for meta. Hope to be more active here :) #twitter"* ***(Twitter)***

𝕏 **(Twitter)** - *"I just want to say [redacted] #HULU for throwing ads into the middle of a documentary you released about 9/11! Immediately canceling my subscription. Unbelievable and so god damn heartless"* ***(Hulu)***

Service: The service sub-category comprises any inclusiveness feedback related to customer support issues with an app. Although many software apps are straightforward to use and require little manual intervention from app developers in everyday usage, occasional situations become awry, and users interact with customer service or support. Essentially, to foster an inclusive user experience, users should have access to reliable customer service that can help fix user problems when they occur. However, I observed a number of complaints from users about how they receive unfair treatment from app developer representatives who cannot or do not want to resolve conflicts. This impacts the user's ability to continue using the app. In particularly challenging circumstances, users find it difficult to contact customer support to report their problems.

💬 **(Play Store)** - *"Venmo just does not seem to work. I have called customer service and they were useless. They have no idea how their product works. I can no longer make purchases and this is made everything difficult. After this experience I would actually recommend people just start deleting this app. It kind of sad because it was useful until the latest update now I can even connect my account. If I could give zero start I would."* ***(Venmo)***

💬❓ **(Reddit)** - *"My personal account got hacked and disable but my business account linked to it still on. I cannot access both. I filled a claim on 10-14 and still nothing! It says due to the few reviewers it may take longer. The thing is my account will be permanently disabled in 8 days if nothing is done. I am thinking about buying an Oculus Quest to try to contact their customer service... Had anyone done this with success? Please post some hope here."* ***(Facebook)***

In the Venmo example, the user experienced a lack of customer service that renders their account, and by extension, the Venmo app, useless for the user. The inability to resolve the user's problem creates a sense of exclusion from Venmo's intended functionality. In the Facebook example, the user faces the potential that their account is permanently disabled. The severe lack of customer support is so dire that the user mentioned that they are considering spending money to buy an additional Facebook product so that they may have access to other customer service.

💬❓ **(Reddit)** - *"I attempted to use my Coinbase Card in a supermarket and it was declined. Upon checking my Coinbase account I receive an error message saying: "You are currently blocked. Sorry, account temporarily disabled. Please contact support"*

The problem is, support have never answered my support case. I used their telephone support, which said to email them... I have never used Coinbase for any nefarious purposes and have no idea whatsoever of the reason why they would lock my account. ..." ***(Coinbase)***

💬 **(Play Store)** - *"Facing a worse scenario and Your Customer service is unacceptable and very weird! My facebook account has been locked and I'm not able to login regarding that I have mailed you through this gmail but not responded yet. Kindly Reply to my query asap."* ***(Facebook)***

𝕏 **(Twitter)** - *"Making me wait for 15 mins and then connecting call after that disconnecting saying network issue. After they call back and asks for otp which I didn't receive ND again disconnects"* ***(Amazon)***

For the user blocked from using Coinbase, they tried calling and emailing, which all failed to connect with any customer service. Similarly, users tried emailing and calling in the other two examples but were unsuccessful. These examples demonstrate how users perceive a lack of fair customer service from app developers regarding their problems. More importantly, these examples show a lack of inclusiveness from app developers to restore services to users currently unable to use their respective apps.

4.1.2 Technology

Technology inclusiveness related user feedback is the second most prevalent category out of the 6 (R: 23.7%, A: 20.69%, X: 32.8%). Technology refers to concerns regarding users experiencing exclusion from a software or feature due to certain technology restrictions enforced by the developer. These limitations arise when developers insist on users having access to particular ***website***, ***device***, or ***network***. Inadvertently, developers may exclude users by specifying certain devices or mandating third-party account access, which may not be applicable to all users. Therefore, user feedback pertaining to the integration of any technological aspect, assuming that users have access to the technology when they do not, falls within this category.

Device: Users often face the obstacle when features for an app no longer work on a particular device.

💬 **(Reddit)** - *"hey, so I use the discord browser on my tablet to manage multiple*

accounts, but today I just seem to get a blank grey screen, and visiting their website there's no longer any way to login or access to web browser. Is anything else having this issue? It's really obnoxious, because I'm even visiting through the desktop website, and the formatting on their web browser has been working fine up until now - it's like they've gone out of their way to check my browser and specifically disable a totally functional feature" ***(Discord)***

💬 **(Play Store)** - *"Videos stutter and play so choppy on my galaxy s22 ultra. This has gone on since I bought this phone. Please fix this app so videos will play smoothly on this phone. Other people with this phone have had the same complaints. PLEASE FIX!!!"* ***(Amazon Prime Video)***

𝕏 **(Twitter)** - *"How is @hulu not supported on the @SamsungMobile Galaxy s22 Ultra?"* ***(Hulu)***

From these examples, I identified that the lack of support for apps or features in various devices is a recurring problem echoed by users across all three of Reddit, Twitter, and Play Store. The impact of inclusiveness related to technology is often quite significant as it is the difference between whether a user can even use an app. As described in the example from Twitter about Hulu on Samsung Galaxy S22 Ultra, the app is not even supported on the aforementioned device.

Network: In addition to devices, I also observe instances of users retelling issues surrounding their network signal.

💬 **(Play Store)** - *"the app never works with my wifi. I have to Initiate the app via cellular data, load the show I want to watch, and THEN initiate my wifi connection. I do not have slow internet."* ***(Hulu)***

I found users discussing problems that manifest after software updates. Organizations frequently update their software or opt to discontinue the product on devices or networks without providing any prior information to users. This generates frustration and a feeling of exclusion from the software amongst users.

𝕏 **(Twitter)** - *"... @Microsoft is suddenly refusing mails from my servers, auto-responding "part of [..] network is on our block list (S3150)". Their postmasters, though, just literally told me the 3rd time, that they "were unable to identify anything on our side". Frustrating."* ***(Microsoft Outlook)***

💬 **(Reddit)** - *"I have an iPhone SE (2020), and I'm on a "pay as you go" phone plan. When I text with Whatsapp, it uses wifi to send my messages without any problems. However, when I make a voice/video phone call, my Whatsapp will use my "pay you as go" minutes instead of wifi. What is the problem? How do I prevent Whatsapp from using my "pay as you go" minutes when making voice/video phone calls with Whatsapp? I never had this problem when using Whatsapp on an Android phone."* ***(Whatsapp)***

In these examples, I discovered the difficulty that users experience when the service provided by the apps suddenly changes without warning. In the complaint from Reddit about WhatsApp, it appears that WhatsApp is only causing this additional surcharge after the users switch their phone from Android to iPhone.

Website: Similarly, I observe users facing restrictions on the basis of single sign-on or app partnerships. Some apps require users to sign up to another website or app first before granting access. In one example, a user is refused the ability to change the email for their account despite them no longer having access to the email. I further observe instances of users not being able to link information that otherwise works on other platforms.

💬 **(Play Store)** - *"It would be wonderful if I could change the email address so I could use my account that already use my phone number so now I'm just ..... cuz I no longer have access to that email NEED TO CHANGE MY EMAIL ADDRESS OR REMOVE MY PHONE NUMBER SO I CAN CREATE MY ACCOUNT"* ***(Afterpay)***

💬 **(Play Store)** - *"Cant even link to my bank account. I used Plaid on another app and no issues. Just venmo having issues."* ***(Venmo)***

💬 **(Play Store)** - *"Ease of using this app is the biggest joke of the millennium. A person gets a new device and tries to get back in to their account and Facebook and it's associated sites won't permit you to log in. Sends you to a security question where their computers make you select statements you made from 3 months ago. Then it determines that it can't identify its you and blocks you from using the site. Hopefully someone else will develop a better version of Facebook so everyone can delete this obsolete app."* ***(Facebook)***

4.1.3 Privacy

The third most common type of inclusiveness user feedback is privacy (R: 11.8%, A: 16.2%, X: 10.7%). This category relates to any user feedback about ***privacy*** and ***security*** concerns such as system permissions, personal data access, and compromises. Personal data access covers a large breadth of areas and includes personal information like banking information, social security numbers, geographical location, and others. Personal data refers to situations where users are limited from using any software due to the exploitation of their account on the app. A hacked account is a common example of this exploitation. Additional user feedback involves concerns about the security of their personal data.

Privacy: Privacy indicates the desire for control over one's information. Through my analysis, I found numerous examples of users complaining about apps requesting users to give up their privacy.

(Reddit) - *"I have been thinking about starting my own depop shop... and at first I thought I was supposed to have a business paypal. In the middle of setting up my account I heard that you needed a personal account instead of a business account. So I went back to my business account and tried to close it, but its making me fill out my ssn and all this stuff about my business when I dont even have one in the first place. I just dont know what to do because I really dont want to put in my ssn."* ***(Paypal)***

(Play Store) - *"To use this app must turn on GPS (location)"* ***(Facebook)***

These user feedback share a common theme. Users frequently cannot use an app or their features without first providing *additional* data. In the case of PayPal, the user seeks to close their account, but they are unable to do so without providing their social security number, which they do not even have. Moreover, in the Facebook example, the user cannot use the app without first turning on GPS and granting the app to the user's location. In either situation, the user cannot complete their desired task without giving up more of their personal data as they face the prospect of being excluded from the software. From the feedback, it is inferred that the apps have not successfully convinced the users that providing additional personal data is warranted for the circumstances.

I also observe privacy concerns regarding the terms of service and privacy policies that apps prescribe. In the following example, a user is frustrated that signing up

for BeReal means agreeing to a policy for user data that gives unlimited rights to BeReal. The app developer would be able to do whatever they want with user data once uploaded. As described earlier in the section, this is a clear example of an inclusiveness related concern as users feel excluded from using an app without, in this particular case, giving *unlimited* rights to their data.

𝕏 (Twitter) - *"@BeReal_App this is crazy. I wouldn't care if you'd just store those BeReals at your servers, but giving you rights to do anything you want with them? No, thanks. I'm deleting the app."* ***(BeReal)***

Security: Security refers to one's need for protection from any harm, in this context, protection of account information. I discovered that hacking and security incidents frequently put users in a challenging situation where they cannot access an app or have a reduced sense of customer trust.

💬 **(Play Store)** - *"Without my permission I'm getting OTP from the what's app like it showing your what's being registered in new device I don't know who was using my number...it's affect the customer trust "* ***(Whatsapp)***

💬 **(Play Store)** - *"Many IT Team hacks my whatsapp contact And I can only block them in WhatsApp, why strict action is not taken against them, why there is no option to write anything against them in the report. (FROUD LOTTERY SCAM, LOAN RECOVERY AGENT HARRASMENT SOFTWARE) LIKE ETC,"* ***(Whatsapp)***

💬 **(Play Store)** - *"now today it suddenly shows me that someone other owner changed my account password how's that even possible I shared my id with my friend nd she said it's fine I am very tensed about this please solve this problem as soon as possible"* ***(Instagram)***

In all of these instances, users report a perceived fraudulent activity surrounding their account or exploitation of their account. Since access to one's personal account is a basic requirement for any app, software app developers should take precautions if account security issues are a recurring concern. The other concern from users is the perceived inaction or lack of options for users to report fraudulent activities.

4.1.4 Demography

The fourth category (R: 12.6%, A: 13.1%, X: 13.9%) amalgamates concerns about user feeling excluded due to their demographic factor including ***age***, ***gender***, ***language***, ***location***, and ***socio-economic status***.

Age: One common sub-category of demography inclusiveness feedback concerns users griping about age-related policy violations.

(Reddit) - *"I hate Twitter so much man Ok so my Twitter got deactivated cause I made it wheN I was 10. I changed my DOB and it suddenly locked me out. This was on Sept 7 and now Im tryna log in and i cant get my account back? Its not recognizing it. Whats going on? How do I get it back? I got 8 years of stuff in there"* ***(Twitter)***

Software organizations must follow the law, and many jurisdictions have stringent laws regarding child protection. In the Twitter example, it seems that the user is concerned because they are locked out of their account, which has close to a decade of content. The user is of age now, but it seems that there is limited guidance for supporting the user to unlock their account. I identified that age inclusive problems are closely related to the service inclusive category. When users have a problem, they tend to try to contact support, but as described in the *Service* sub-category (Section 4.1.1), users complain that support is unresponsive or negligent. In the example below, a user attempts to contact TikTok to resolve their age dispute, and they never heard back from the app neither did their age verification get approved.

(Play Store) - *"The development team never gets back to you. I had submitted a claim to get my age verification done after not being able to have the option to add my birthday even though I did it when signed up. Reported it. Team followed up. I then did as they asked. I submitted a second claim and got no response. They ghosted me. It's absolutely ridiculous I followed instructions only to never hear back. Can't view age restricted videos but im over 18. I have now told everyone to avoid this app."* ***(Tiktok)***

Location: Another common sub-category is location, which encapsulates any user feedback about exclusion from using an app based on a user's location. Specifically, location may refer to a user's country, city, province, state, or anything to do with

the geographic coordinates.

💬 **(Play Store)** - *"Why ain't Reels available for everyone in every country and Instagram music? It sucks not being able to hear the sounds some [people] post just because IG music isn't available in my region."* ***(Instagram)***

💬 **(Reddit)** - *"My account got locked and it asked me to verify my phone and email but my country code isn't even an option how am I supposed to verify? Country code is South Africa (+27) "* ***(Twitter)***

💬 **(Reddit)** - *"I am aware that paypal supposedly has a minimum amount you can withdraw (around a dollar) but in new zealand you cant withdraw any amount of money. Any amount of money i try to withdraw (tried 20$) will say "does not meet the minimum amount required to withdraw". I now have money sitting there forever, and need to do some other sales but cant if i cant withdraw the money. On top of this, paypal (that ive been able to see) only does support via phone, on numbers that you cannot call from new zealand. So i essentially have no contact for support, and hence why im asking for help here."* ***(Paypal)***

In these examples, users complain about a variety of problems that all relate to exclusion from using the apps properly due to restrictions based on their location. For Instagram, the user feels excluded from the app as music is not available in every country, and the user's country is not on the list of exempt countries. Similarly, a user of Paypal in New Zealand laments that they cannot withdraw money due to location restrictions. In addition to this, the PayPal user cannot request support either, as PayPal does not provide a local phone number for people in New Zealand.

💬 **(Reddit)** - *"Hello! I am trying to open an Etsy shop, but when I try to set up my shop security, by the authenticator app, I just can't... Every code I entered was a correct one, but Etsy said they weren't valid. Now I have too many failed attempts and do not know how to proceed... The other methods do not work for my country... Any tips? Has anybody dealt with this before? Thank you..."* ***(Etsy)***

💬 **(Play Store)** - *"I can't close my account after sending funds across the ocean because it's up to a random person in another country to have the knowledge to accept them once sent because of a redundant verification system."* ***(Paypal)***

I also observed similar location problems regarding account opening or closing.

An Etsy user complains that they cannot open an Etsy shop because they cannot use codes from the authenticator app. However, this was their last opportunity after exhausting all other available attempts to open the account.

💬 **(Reddit)** - *"Trying to buy with N26 card from France Used to buy without any problem, now I get this: "We currently do not support the bank cards issued in your country, but we are working on making this a possibility as soon as we can." Did they ban Germany from Binance? What's happening?"* ***(Binance)***

I also found location issues that occur from sudden shifts in geographic restrictions. A feature that used to work for users is modified and no longer works for users from a country or state.

𝕏 **(Twitter)** - *"Pluto TV as Xumo on Firestick needs a VPN in UK"* ***(Pluto TV)***

💬 **(Reddit)** - *"Search feature is difficult, many items are not found in intuitive categories, searches bring up 1000 "related" or "suggested" items and you must search thru pages of listings sometimes to find the item you are looking for. When you choose an item and add it to your cart, the app will not allow you to change the delivery method that was chosen by default when you opened the app. The app does not save my location and the "use my location" feature does not work. My default is always Sacramento CA? "* ***(Walmart)***

Language: Closely related to location, I found an assortment of user feedback about language issues. Users often complain about a perceived lack of inclusiveness, where their choice of language is not supported by an app.

💬 **(Play Store)** - *"Unusable. Can't deposit money. Can't trade. No proper UI in English."* ***(Binance)***

💬 **(Play Store)** - *"Usually good, but the change in notification is annoying and confusing, a lot of creators are harassed and receive strikes when they are being harassed, the language configuration is not able to detect the language in content and often send me content in languages I signaled as not interested. ..."* ***(TikTok)***

These examples demonstrate that criticism is directed at the choice of languages in the app and that their choice of content is excluded from the app. End users cannot use an app if the app developers do not support their preferred language of choice.

Gender: I also found user feedback about gender inclusiveness, albeit to a limited extent. In the following example,

💬 **(Play Store)** - *"This is such a useful app and it has everything, literally. It has clothes of every style and it's just overall a great app with also things for your house. Everything about this app is great, the only thing I would say could be improved is that the app is based around women and mainly women and maybe could have more advertising for men."* ***(SHEIN)***

Socio-economic status: Another common problem amongst frustrating users is the limitation caused by a user's socio-economic status and payment preference. A number of users explain the issues related to apps being too expensive for them to continue using. This scenario illustrates how economic status impacts a user's access to the app. When a user is unable or unwilling to pay for premium subscription services, the user experiences a sense of exclusion despite their fondness for the app.

💬 **(Play Store)** - *"Premium features are too expensive. Video/audio calls are very poor when one end has a slow device or poor connection."* ***(Telegram)***

💬 **(Play Store)** - *"This is a very good place to watch anything you want!! But the price is pretty expensive.....I would really love to keep watching my precious anime on here but I can't keep up with the payments"* ***(Hulu)***

4.1.5 Usability

This category covers any user feedback regarding concerns about software usability (R: 15.3%, A: 8.0%, X: 11.5%). In particular, usability focuses on ***visual*** and ***sound*** aspects of accessibility and preference for user interface design.

Visual: Visual usability concerns are more widespread than audio. There are numerous incidents of users feeling some sort of visual inhibition that prevents them from a painless experience.

💬 **(Reddit)** - *"Cant use the app cause the screen is too small to reach [the] button at the bottom wtf who Programms something like that?"* ***(TikTok)***

💬 **(Reddit)** - *"Upgraded to new version 2.2237.5 desktop (windows) and it's horrible. Font size is tiny and no way to change? Messages are centered and not left-*

justified like before. Unable to quick reply - now have to click a menu first, and select "reply" ? Put this back the way it was or off to SIgnal et, al. I go" ***(Whatsapp)***

𝕏 (Twitter) - *"The UI of the #YouTubeMusic on #Roku is soooooo annoying :("* ***(Roku)***

(Reddit) - *"Is it me or has Instagram WEB changed their ad-blocker policy? I use instagram Web occasionally and without my ublock origin being disabled the main features/buttons are not able to be clicked and not visible either. It couldn't have been more than a few weeks since I could use Instagram with adblocker ON and fully functionally. Now I am getting Ads every other scroll and it's unbearable for an autist like me... I can't tolerate the format and frequency of those ads so I'm going to have to stop using it if no solution can be found"* ***(Instagram)***

In all of these examples, users report their problems experiencing difficulty using the app due to the visual usability, either due to the small font or design that may not consider the users' device. It is representative of a lack of inclusiveness in the user interface design. To offset these issues, app developers would need to consider designing a user interface that considers these diverse user needs. The WhatsApp example demonstrates that the font size in the user interface is small for the user, but they do not have an option to modify it. The Tiktok example illustrates a problem when the screen is too small, and the app does not consider this problem. To support diverse users with different visual requirements, organizations should consider implementing various visual options.

Audio: In addition to visuals, I found audio related usability inclusiveness in user feedback. As suggested by the sub-category name, audio concerns user feedback where the user describes sound issues.

(Reddit) - *"Discord makes every mic I use super Quiet I've tried basically everything in terms of my discord and my windows settings trying to make my mic sound louder, but nothing works. My friends can barely hear me when I'm on 200%. I've had this issue for a while and it's really frustrating. ... "* ***(Discord)***

(Play Store) - *"Voice call not always audible enough."* ***(Whatsapp)***

I discovered instances of users lamenting the lack of audio quality in the usage of the apps. The basic function of communication apps is defeated when users cannot

make basic voice calls, and the microphone audio fails to work.

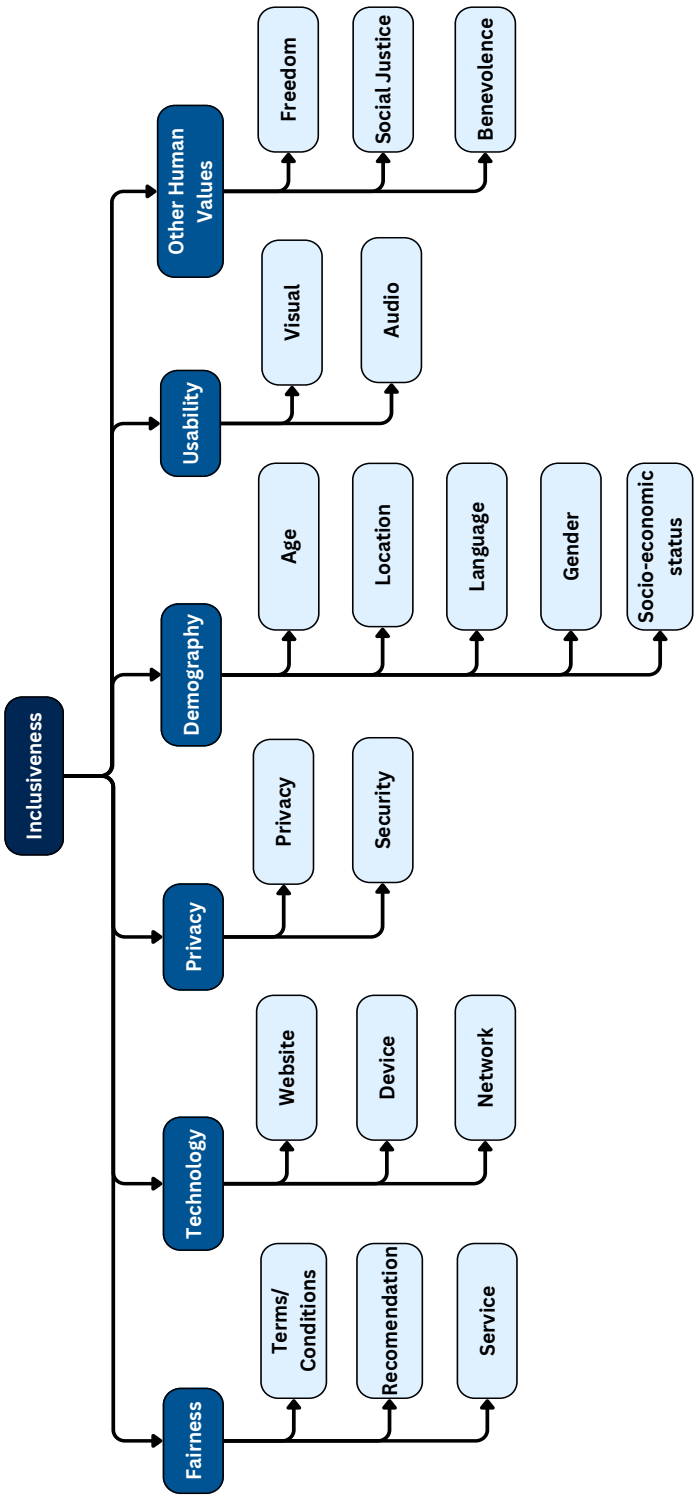

Figure 4.1: Taxonomy for inclusiveness related user feedback

4.1.6 Other Human Values

I found various posts that indicated the violation of different basic human values and resulted in the user feeling excluded from the software. Basic human values refer to "principles that guide social life and are modes of conduct that a person likes or chooses among different situations" [34]. To better structure and present such related user feedback, I draw upon Shalom H. Schwartz's [39] theory of basic human values. The theory amalgamates 58 values grouped into 10 categories. In my dataset, I identified the presence of 3 values ***freedom***, ***social justice***, and ***benevolence***. I found users expressing exclusion from the app when they perceive a violation of these human values.

Freedom: This sub-category particularly advocates for "freedom of thought or speech," often tied to ideals of equality. I identified users describing their feelings of exclusion from freely providing their thoughts or interjecting their voices. Many users express frustration about being restricted from voicing their opinion in any given app.

💬 **(Play Store)** - *"They ban users for political views. Not acceptable!"* ***(Cash App)***

💬 **(Play Store)** - *"It restricted our account when We say true words and we say true words about human rights because if we say about some places or people that don care about human rights when we say our account restricted"* ***(Facebook)***

𝕏 **(Twitter)** - *"You officially lost a customer you are a big time loser facebook learn to respect freedom of speech!"* ***(Instagram)***

💬 **(Reddit)** - *"I got banned from Tik Tok because I posted a picture of Turkish Homelander, literal censorship "* ***(TikTok)***

💬 **(Play Store)** - *"I remember when Facebook was fun, not it's not any longer. It was a place to connect with family and friends. A place to talk and share thoughts and ideas but since it started to censor speech it's no longer a place to be."* ***(Facebook)***

All these examples suggest that users with specific opinions may feel marginalized or excluded from these platforms as their ability to freely express themselves is being restricted.

Social Justice: I found many user complaints regarding social justice, leading users to experience a sense of exclusion within the app. Social justice refers to a commitment to ensure that all members of society, irrespective of their race, religion, gender, or other characteristics, have equal access to opportunities and resources, in this context, equal access to the app and all the features. However, my results indicate a different scenario.

💬 **(Play Store)** - *"This app is based on Islamophobia. We cannot post openly about Islam. If we do, they give restrictions on our accounts much needed to improve"* ***(Facebook)***

💬 **(Play Store)** - *"They support racism and conspire in front of the Arabs. ..."* ***(Facebook)***

💬 **(Play Store)** - *"Completely biased platform always support muslims and Christians. Always hindu phobic"* ***(Instagram)***

Benevolence: This sub-category pertains to users feeling excluded from the software due to their concern about the well-being of people they interact with on a regular basis, i.e., family and friends [39]. I particularly found instances where users believe the software lacks child-friendly content and lacks features that could improve the overall family experience. The absence of such consideration leads the user to perceive that the software lacks inclusiveness.

💬 **(Reddit)** - *"My young kids love watching YouTube. But I'm really uncomfortable letting them have unfettered access to the full YouTube content. There is a lot of disturbing content and the algorithms are known for luring people into extreme content. I have Roku TVs and the YouTube for Kids app is unavailable. ..."* ***(Youtube)***

💬 **(Play Store)** - *"... The recent video released of Disney internal meets proves they dont have the best interest of children in mind. When you decide to make wholesome content again and stop pushing sexuality on children is when we will spend money with you again. From now on our home will be Disney free. ..."* ***(Disney+)***

💬 **(Reddit)** - *"Trying to add grandma to my daughter's kid messanger and keep getting an error ... when I try to approve her and nothing has fixed it. ... We live far away from my parents so this is really the only way my mom gets to see her and the fact that it's not working only for her is very frustrating... "* ***(Facebook)***

> **RQ1:** I developed a taxonomy of inclusiveness following a STGT approach using user feedback from Reddit, Google Play Store, and Twitter to structure the different types of inclusiveness related user feedback. I found six major categories of inclusiveness, ranked in order of prevalence: *fairness*, *technology*, *privacy*, *demography*, *usability*, and *other human values*.

4.2 Inclusiveness in Different Types of Apps and Source of Feedback

This section answers the second research question: *How does inclusiveness related user feedback differ for different types of apps and sources of feedback?*

In this thesis, the dataset of 50 apps stems from 5 different domains of software: business, entertainment, financial, e-commerce, and social media. Each software domain contains ten apps. The qualitative analysis presented in Section 3.2 enabled me to identify the prevalence of the six major categories from the taxonomy (sec. 4.1). In addition, the analysis allowed me to uncover the distribution of the categories in each domain. Table 4.1 illustrates the total number of inclusiveness related user feedback found in the five domains of apps across the three sources. I found that 633 out of 1,211 of the inclusiveness related user feedback emerged from *social media* apps. This is followed by *financial* apps (212) and *entertainment* apps (196). Whereas *e-commerce* (99) and *business* (71) software contain the least number of inclusiveness posts.

The distribution of the inclusiveness categories is shown in Figure 4.2. *Fairness*-related concerns represent the most frequently discussed (355 out of 1,211) category and is more prominent for *social media*, *financial*, and *e-commerce* apps. On the contrary, *technology* is the more dominant category in *business* and *entertainment* software and a close second for *social media* apps. *Social media* consists of a significant number of user feedback from the 6 categories, particularly for *fairness, technology, and usability.* These three from *social media* surpass even the most popular categories for the other app types. However, the other three categories, *privacy*, *demography*, and *other human values*, are still more frequently occurring than in the other apps.

I further observed varying levels of popularity for each category depending on the type of app. For example, 45 out of 121 inclusiveness concerns for *financial*

Table 4.1: Total number of inclusiveness related user feedback in the 5 types of apps from Reddit, Google Play Store, and Twitter. (Here the columns indicate the following: Tech. is Technology, Priv. is Privacy, Fair. is Fairness, Demo. is Demography, Val. is Other Human Values, and Usa. is Usability)

App Type	Source	Tech.	Priv.	Fair.	Demo.	Val.	Usa.	Total
Business	Reddit	20	3	4	3	2	5	37
	Play Store	6	3	7	1	0	2	19
	Twitter	8	1	0	0	1	5	15
	Total	34	7	11	4	3	12	**71**
Entertainment	Reddit	34	6	18	22	14	25	119
	Play Store	13	1	12	9	3	5	43
	Twitter	8	1	3	13	3	6	34
	Total	55	8	33	44	20	36	**196**
Financial	Reddit	17	35	65	26	4	3	150
	Play Store	9	8	30	5	3	1	56
	Twitter	3	2	1	0	0	0	6
	Total	29	**45**	96	31	7	4	**212**
E-commerce	Reddit	5	5	20	15	1	1	47
	Play Store	4	1	22	9	1	1	38
	Twitter	2	2	8	1	1	0	14
	Total	11	8	50	25	3	2	**99**
Social Media	Reddit	93	35	103	24	29	75	359
	Play Store	46	48	55	25	26	21	221
	Twitter	19	7	7	3	14	3	53
	Total	158	90	165	52	69	99	**633**
Total	Reddit	169	84	210	90	50	109	712
	Play Store	78	61	126	49	33	30	377
	Twitter	40	13	19	17	19	14	122
	Total	287 (24%)	158 (13%)	**355** (29%)	156 (13%)	102 (8%)	153 (13%)	**1211**

apps are about *privacy*, which is to be expected as *financial* apps obtain confidential information from users, and it is their responsibility to adequately safeguard this

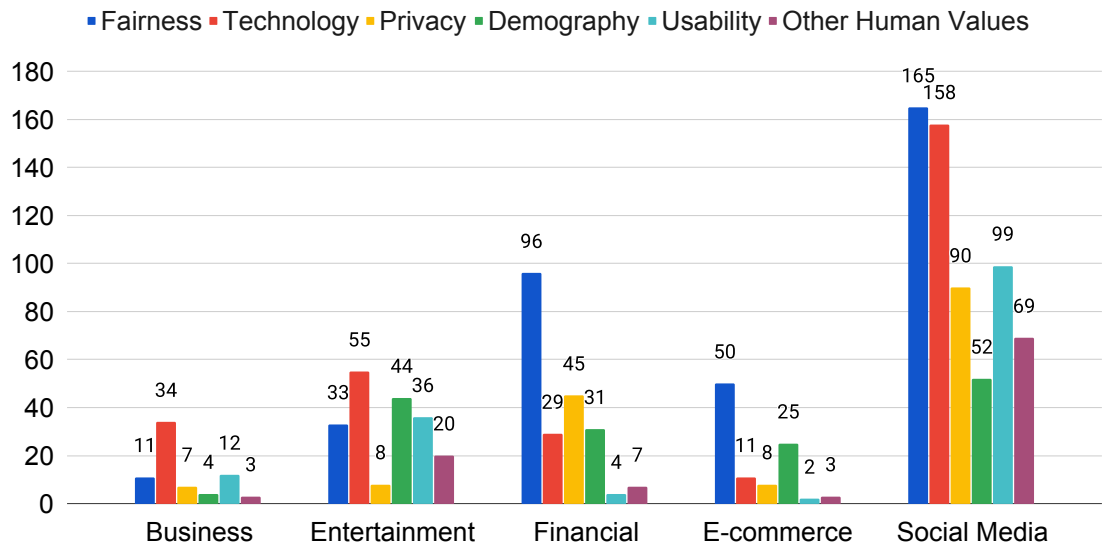

Figure 4.2: Category distribution across five types of apps

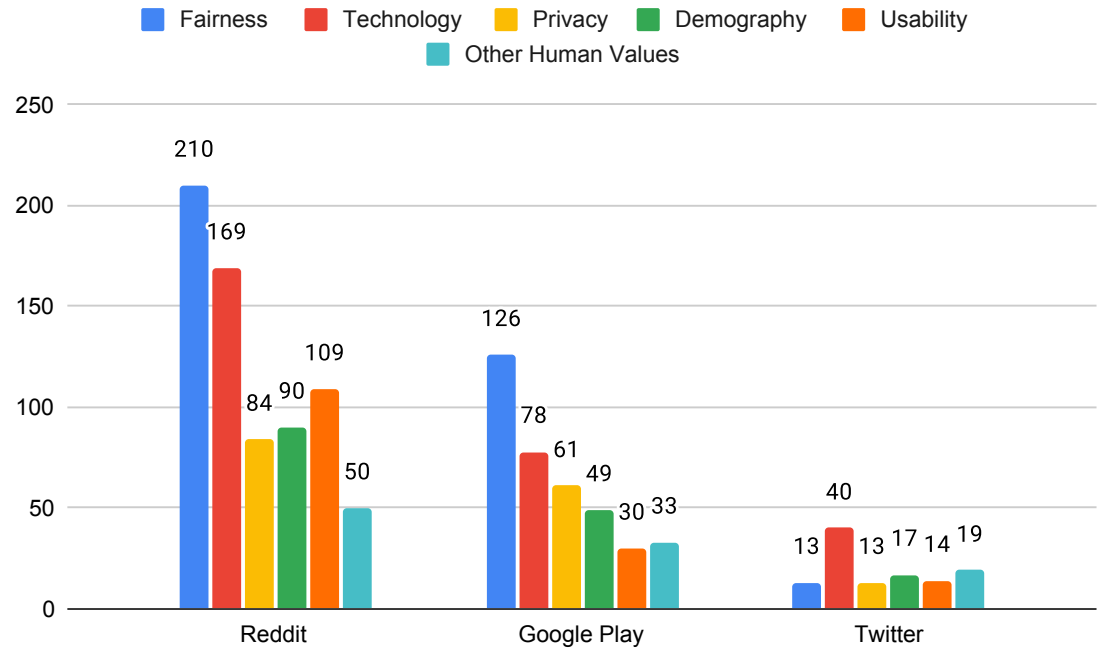

Figure 4.3: Category distribution across three sources

data. Similarly, I found *demography* to be a popular category for *entertainment* type of software. *Demography* being popular in this context appears reasonable, given that location, language, and socioeconomic status may impact the content provided

or suggested in *entertainment* apps. For *business* apps, I found *technology* as the most popular. Since productivity is critical for *business* apps, technological restrictions that prevent users from completing their tasks and work would definitely be a high area of concern.

Moreover, figure 4.3 and table 4.1 show that Reddit has a far greater number of inclusiveness user feedback in comparison to the other two sources. After all, *Reddit* has by far the largest character limit among the sources and is a platform for users to engage in long form discussion. Therefore, users often share a breadth of details, including inclusiveness problems. Particularly, I found the inclusiveness feedback materializing in the *fairness* and *usability* categories, where users are likely to give detailed descriptions of their unfair or poor usability experience.

In contrast, I observe slightly different popular categories for *Play Store*. I discovered that *privacy* is a more prominent category for *Play Store*, which makes sense as users who encounter *privacy* or *security* issues in an app can directly voice their reviews to warn other potential app users. For *Twitter*, the popular categories differ slightly, and *technology* is by far the common category. Since *Twitter* is most often a platform for users to contact support or voice real time feedback, users will discuss device or network *technology* problems when they occur.

> RQ2: Fairness inclusiveness is most prominent for social media, financial, and e-commerce apps, whereas technology inclusiveness is most common for business and entertainment apps. Reddit contains more inclusiveness feedback in comparison to Twitter and Google Play. Fairness is most common for Reddit and Play Store, but technology is most common for Twitter.

4.3 Automated Classification

This final section reports on the third research question: *Can we automatically identify inclusiveness related user feedback from online sources?*

For the final step, I assessed the effectiveness of the five classification models, as detailed in Section 3.3 through evaluating on the same dataset. I fine-tuned the pre-trained models on each of the sources and measured the performance in terms of precision, recall, F1-score, and accuracy. The performance results of the five models for each source are outlined in Table 4.2. When reporting the best performing

classifier, I use the F1 score as it helps to seek a balance between precision and recall.

Table 4.2: Results from five deep learning models on classifying between inclusiveness and non-inclusiveness.

Source	Model	Precision	Recall	F1-Score	Acc.
Reddit	GPT-2	0.903	0.812	**0.838**	**0.813**
	BERT	0.895	**0.813**	0.837	0.803
	RoBERTa	0.899	0.785	0.802	0.772
	XLM-RoBERTa	0.887	0.768	0.802	0.772
	BART	**0.904**	0.769	0.805	0.776
Google Play Store	GPT-2	0.934	0.775	0.828	0.778
	BERT	0.930	**0.806**	**0.849**	**0.811**
	RoBERTa	**0.940**	0.739	0.802	0.737
	XLM-RoBERTa	0.926	0.801	0.845	0.798
	BART	0.928	0.767	0.822	0.766
Twitter	GPT-2	**0.992**	0.851	0.911	0.851
	BERT	**0.992**	0.853	0.913	0.852
	RoBERTa	**0.992**	0.863	0.919	0.861
	XLM-RoBERTa	0.991	0.850	0.911	0.851
	BART	**0.992**	**0.884**	**0.930**	**0.881**

The overall evaluation results are best for Twitter user feedback. All the evaluation metrics are above 0.85 for Twitter for all five different classifiers. Among the five classifiers, I found BART to have the best results in terms of F1-score, achieving a value of 0.930. One possible reason for the better performance on Twitter may be that Twitter data has a lot of unrelated discussion (e.g. ads), making it relatively easier for the classifier to differentiate between inclusiveness and non-inclusiveness. Since there are many Twitter posts that consist of completely random dialogue or topics obviously unrelated to inclusiveness, this may help the models classify the user feedback.

I observed that the classification results for app reviews are not as good as Twitter's (i.e., roughly 8-12% lower). For Play Store, the best performing classifier is BERT, with an F1-score of 0.849. There are many user feedback reviews that report

bugs concerning the apps, but not all of these bug reports are about inclusiveness. This may be a reason for the increased difficulty for the classifiers to identify inclusiveness in app reviews.

Finally, I found that the classifiers exhibit a comparatively lower performance on Reddit when compared to that of other feedback sources. The best performing classifier for Reddit data is GPT-2 (F1-score: 0.838), whereas the overall performance for classifiers for Reddit is roughly 8-14% lower than Twitter. The most likely reason why performance on Reddit is lower is that user feedback from Reddit is often complex and detailed in nature, which may contain an assortment of topic discussions. Identifying the inclusiveness aspect in a Reddit post is less clear cut than in a Twitter post or app review.

> RQ3: Employing deep learning models for classification proves quite effective in identifying inclusiveness-related user feedback. GPT-2 demonstrates the best results for Reddit, BERT excels with app reviews, and BART achieves the best results for Twitter.

Chapter 5

Discussion

In this thesis, I proposed a revised taxonomy for inclusiveness related user feedback derived using a socio-technical grounded theory approach [18] conducted on user feedback from fifty popular software apps. The approach involved labelling over 23 thousand user feedback posts across three popular sources of user feedback: Reddit, Google Play Store, and Twitter. The analysis resulted in 1,211 user feedback related to inclusiveness. Using the labelled set, I present a machine learning model that automatically classifies the user feedback from the three sources into inclusiveness or not. The best F1 scores for Reddit, Google Play Store, and Twitter are respectively 0.838. 0.849, and 0.930.

5.1 Comparison with Existing Taxonomy

1, 211 inclusiveness related user feedback across the various platforms is a sizable amount that organizations should not neglect. Moreover, neglecting the inclusiveness related needs can lead to a negative perception of the software among diverse end-user groups. This may result in the users feeling excluded, in retrospect, opting for a different app. Therefore, companies need to consider the inclusiveness concerns to ensure they do not unknowingly overlook the inclusiveness aspects.

Previous work on human aspects provides a preliminary introduction to inclusiveness based on their analysis from GitHub and Google Play Store [20] from both developer and user perspectives. However, my research focuses specifically on inclusiveness from the end user perspective. Furthermore, the prior work focused on open source software, which may not reflect the entirety of end users nor software

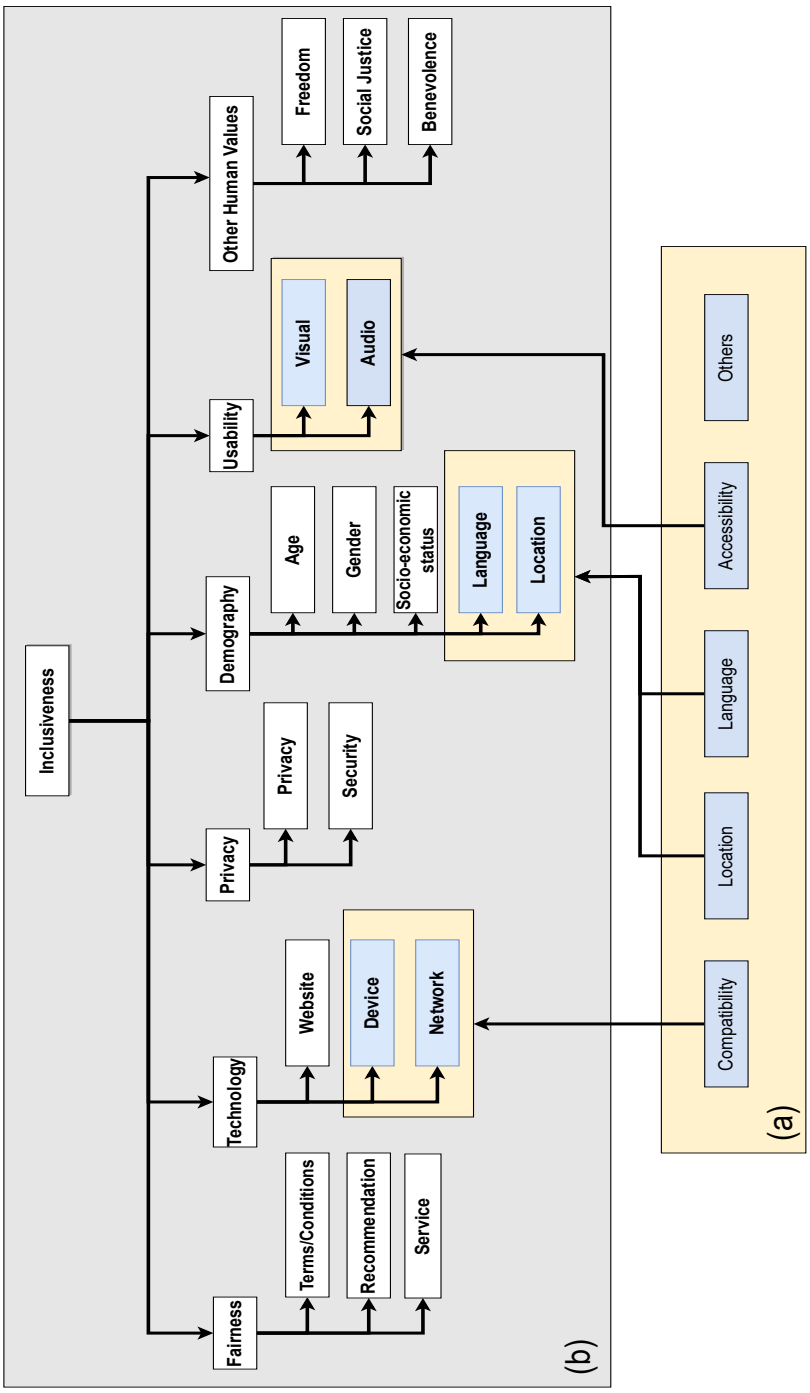

Figure 5.1: Comparison of revised inclusiveness taxonomy with previous taxonomy from [20]. (a) indicates the previous taxonomy, and (b) indicates the revised taxonomy. Highlighted in yellow are the similar sub-categories that are incorporated in the revised version.

apps. Only a fraction of users use open-source software, and even fewer provide user feedback. The prior research reported only 31 posts from Google Play Store pertained to inclusiveness, whereas my research identifies a higher amount (1,211), indicating a better representation of inclusiveness. Additionally, the vast majority of apps, especially the popular ones, are not open source, ranging from social media, entertainment, and business apps. Thus, an understanding of inclusiveness issues from a larger user base is required. Therefore, I focus on popular for-profit apps used by millions of users across the world to obtain a more diverse representation of users. My research shows that all three sources, Reddit, Google Play Store, and Twitter, provide inclusiveness related user feedback, which can be leveraged to inform organizations about inclusiveness issues.

In the work by Khalajzadeh *et. al* [20], the authors included 5 sub-categories under inclusiveness: compatibility, location, language, accessibility, and others. In this research, I found the presence of these sub-categories, albeit in a relatively less dominant way. I compare the difference between my categories and their sub-categories in Figure 5.1. The compatibility category, for instance, aligns closely with my proposed technology category. Their category is primarily about compatibility across different devices, operating systems, and platforms, considering socio-economic factors. Whereas I focus on a broader exclusion of users due to any technology-related restrictions set by developers. Therefore, I integrated the compatibility category into the broader technology category. At the same time, I include any socio-economic concerns under demography in my proposed taxonomy, as the feedback is more in line with the dimensions of demographics.

Regarding the location and language sub-categories, I chose to reclassify them within the demography category in this research. This was done because the two categories were not as dominant in my dataset compared to the previous work. Hence, amalgamating them into the demography category helped to achieve more clarity and better structure. Furthermore, I encountered only a small amount of user feedback explicitly emphasizing accessibility. However, I observed a broader distribution of posts focusing on visual and audio usability. Therefore, any accessibility posts were subsequently placed under the two types of usability, ensuring a more structured distribution of user feedback related to accessibility concerns.

5.2 Algorithmic Bias: A Barrier to Inclusiveness in Software

My research brings additional concrete evidence of the possible impact that algorithmic bias can have on inclusiveness. The thorough analysis allowed me to uncover various inclusiveness issues that users face while using the software. One of the key problems that emerged throughout the analysis was algorithmic bias. With the rapid evolution of AI, companies are increasingly inclined towards integrating algorithms in their decision making process. As such, many features are automated and often exhibit biases towards certain users [5]. Algorithmic bias predominately originates from underrepresented data and biased methods [3]. These biases create both a perceived and real non-inclusive experience for users.

Earlier in the fairness category, I described how the automated decision-making process leads to the exclusion of users from software or specific features in software. For example, Facebook heavily relies on its algorithms for the content review process. AI decides if a content is allowed on their platform based on community standards [1]. Due to the algorithmic decision making approach employed by Facebook, users frequently complained about being banned from the platform due to policy violations without warning or prior notification. In these cases, users have no idea what action even triggered the ban or account restriction. Despite the company asserting that users can seek a repeal if they believe their content aligns with the community standards, the fairness category revealed a contrasting reality.

💬 **(Reddit)** - *"I wanted to get started with Facebook ads and was really motivated about it. Only to find out that somehow my account on Meta Business was disabled, I tried to appeal but it somehow got rejected. I ended up purchasing another domain in order to create another account but during the creation of the facebook account (even though all the information was different compared to my main account), it instantly got disabled and told me to appeal. I feel like the situation is going to repeat itself. I've tried hard to find some support, but after hours of searching I didn't find anything."* ***(Facebook)***

The fairness category has numerous similar examples where users have a lack of support and must rely on automated algorithms to make further decisions. Similar to this example, many software organizations incorporate AI to make decisions, and

that creates more frustration and a feeling of exclusion amongst the users. A likewise pattern is observed manifesting in other inclusiveness categories as well, such as other human values, where users report frustrations from apps enforcing their beliefs on users. In an ideal scenario, recommender systems would learn from user preferences and patterns and tailor them. However, an opposite instance is observed where users receive engagements and recommendations that deviate from what they actually anticipate or prefer.

In recent years, there have been numerous instances of biased AI systems that have come to light. Famous incidents exposing biased AI systems include the COMPAS recidivism algorithm [22], which had a significantly greater likelihood of incorrectly judging black defendants in comparison to white defendants, whereby black defendants were more likely flagged as high risk. Recently, Meta, formerly known as Facebook, agreed to a settlement after it was revealed they implemented features in its advertising to exclude specific groups of people [43].

Stemming from these examples, the study of reducing bias in machine learning systems is actually a large subject area [7]. Several studies explore reducing bias in AI systems, particularly for those that conduct automated decision making [21, 44]. These studies attempt to research how to minimize algorithmic bias from a data collection, model training, and testing level. However, in this thesis, I indicate that organizations should avoid completely relying on AI for decision making, whether it is adhering to community standards or generating recommendations. Such heavy dependence generates a sense of exclusion among users rather than improving their experience.

5.3 Implications for Practitioners

In this thesis, I present the importance of considering user feedback for inclusiveness, which highlights important implications for practitioners. Practitioners can use the approach exhibited in this thesis to identify inclusiveness related user feedback for their software. Particularly, they can use the taxonomy of inclusiveness to categorize user feedback so that the issues are resolved. Developers are predominantly male, technically skilled, and affluent, which differ significantly from the diverse end-users they serve. Awareness of the inclusiveness issues will allow them to learn and consider the diverse user needs and develop more inclusive software. The taxonomy provided in this thesis is a step towards enabling companies and practitioners to capture the

diverse needs of the users.

Second, my research also uncovers some inclusiveness issues that practitioners should recognize. Even though automated AI systems are useful and are increasingly used, my work shows instances of users experiencing a lack of inclusiveness from these automated systems.

Third, software companies follow a fast paced CI/CD (continuous integration and deployment) approach and often lack the resources to address every single need. My research represents categories identified from five types of software: business, entertainment, financial, e-commerce, and social media. I indicate which categories are more prevalent in each type of software in Section 4.2. Based on the findings, companies can prioritize the specific inclusiveness concerns accordingly.

Fourth, my research proposes an automated approach to identify the inclusiveness related user feedback. Online sources like Reddit, Google Play Store, and Twitter generate user reviews in large quantities. Manually identifying issues, particularly inclusiveness ones, is a challenging task, and companies lack the resources to accomplish this. My approach will enable companies to easily generate inclusiveness concerns from a large amount of user feedback. As such, they can promptly prioritize addressing the concerns.

Finally, in the taxonomy of inclusiveness, I present the end users' perspective, which may conflict with various ethical concerns such as discrimination or hate speech towards a group of people. Such posts can potentially generate a sense of exclusion among certain user groups, undermining a respectful environment. Therefore, companies should carefully select which concerns are worth consideration before incorporating them into their software.

5.4 Implications for Researchers

There are several implications for future research. First, more research should be conducted with practitioners to understand how they address these inclusiveness user feedback, particularly how organizations manage inclusiveness requirements. Second, researchers should pay more attention to studying additional sources of user feedback. There are numerous platforms for users to express their concerns, though I studied three popular sources. These additional sources may help refine the categories and sub-categories in the taxonomy. Further research can be conducted on each domain of app to understand the inclusiveness aspect better. Third, my research presents a large

number of manually labelled inclusiveness related user feedback. Future researchers can leverage this labelled data to conduct further studies on improving the automated classification approach and automated identification of the categories of inclusiveness.

5.5 Threats to Validity

Every research comes with its own set of threats, and in this thesis, I rely on the total quality framework of Roller [37] and their four aspects: credibility, analyzability, transparency, and usefulness to identify the threats and mitigation for this research.

For *credibility*, this research may have the threat of sampling bias because I collected user feedback from 50 apps from the sources of feedback. However, I select apps that have a diverse group of users, and the feedback sources are also common platforms that users often use to discuss concerns. The research also uses standard web scraping libraries to collect the data. Additionally, I try limiting bias by creating a randomly sampled batch of user feedback to conduct manual annotation. I did not seek to give more weight to any particular app or source of feedback.

For *analyzability*, the data was analyzed with two researchers who followed a social technical grounded theory approach [18], where open coding, constant comparison, and memoing are used to analyze the feedback for inclusiveness. Furthermore, the annotators were in constant dialog during the coding process to ensure consistency and remove bias. Since this research leverages user feedback from three popular sources (i.e., Reddit, Twitter, and Google Play) and different apps, I was able to triangulate the analysis with the different sources.

For *transparency*, I provide extensive and rich descriptions of the research methodology. I also provide extensive quotes to support the taxonomy. The entirety of the data is provided in a replication package, including the manually labelled dataset, which may help future researchers.

For *usefulness*, my research aims to shed more insight into the role of inclusiveness in user feedback. More importantly, my research aims to advance the state of knowledge of inclusiveness by providing a taxonomy for the different types of inclusiveness related discussions. In particular, my work encompasses a significant number of user feedback and includes more empirical insights for organizations. I acknowledge that the results may not hold for every software app, but I believe organizations can benefit from the inclusiveness categories as they try to consider the concerns from diverse end users.

Chapter 6

Conclusions

This thesis adopts a socio-technical grounded theory approach to delve into inclusiveness related user feedback. Through in depth manual analysis of over 23,000 user feedback posts sourced from Reddit, Twitter, and the Google Play Store regarding 50 popular for-profit apps, I developed a taxonomy of inclusiveness from the end user perspective. This taxonomy encompasses six primary categories: fairness, technology, privacy, demography, usability, and other human values. The classifier I reported based on the labelled dataset shows that it is possible to automatically identify inclusiveness-related user feedback among general user feedback.

The results indicate to practitioners that these online sources contain a rich collection of inclusiveness feedback that organizations should consider to build more inclusive software products for diverse end users. Furthermore, I present a substantial labelled dataset that researchers can leverage to refine tooling, thereby offering better support to practitioners. The dataset contains 712 inclusiveness user feedback from Reddit, 377 from Google Play Store, and 122 from Twitter. My findings offer a foundation for future research to further investigate each category and improve the automated approach, ultimately contributing to creating more inclusive software.

Bibliography

[1] How does Facebook use artificial intelligence to moderate content? | Facebook Help Center.

[2] K. B. M. . Designing inclusive software in Windows - Windows apps. https://learn.microsoft.com/en-us/windows/apps/design/accessibility/designing-inclusive-software, may 13 2022.

[3] Shahriar Akter, Grace McCarthy, Shahriar Sajib, Katina Michael, Yogesh K Dwivedi, John D'Ambra, and Kathy Ning Shen. Algorithmic bias in data-driven innovation in the age of ai, 2021.

[4] Abdulaziz Alshayban, Iftekhar Ahmed, and Sam Malek. Accessibility issues in android apps: state of affairs, sentiments, and ways forward. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, pages 1323–1334, 2020.

[5] Yuriy Brun and Alexandra Meliou. Software fairness. In *Proceedings of the 2018 26th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering*, pages 754–759, 2018.

[6] Margaret Burnett, Simone Stumpf, Jamie Macbeth, Stephann Makri, Laura Beckwith, Irwin Kwan, Anicia Peters, and William Jernigan. Gendermag: A method for evaluating software's gender inclusiveness. *Interacting with Computers*, 28(6):760–787, 2016.

[7] Joymallya Chakraborty, Suvodeep Majumder, and Tim Menzies. Bias in machine learning software: Why? how? what to do? In *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 429–440, 2021.

[8] Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. Unsupervised cross-lingual representation learning at scale. *arXiv preprint arXiv:1911.02116*, 2019.

[9] Sara Donetto, Paola Pierri, Vicki Tsianakas, and Glenn Robert. Experience-based co-design and healthcare improvement: realizing participatory design in the public sector. *The Design Journal*, 18(2):227–248, 2015.

[10] 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*, pages 38–49, 2022.

[11] The Gaurdian. Twitter apologises for 'racist' image-cropping algorithm. https://www.theguardian.com/technology/2020/sep/21/twitter-apologises-for-racist-image-cropping-algorithm, 2020.

[12] Eduard C Groen, Joerg Doerr, and Sebastian Adam. Towards crowd-based requirements engineering a research preview. In *Requirements Engineering: Foundation for Software Quality: 21st International Working Conference, REFSQ 2015, Essen, Germany, March 23-26, 2015. Proceedings 21*, pages 247–253. Springer, 2015.

[13] Eduard C Groen, Norbert Seyff, Raian Ali, Fabiano Dalpiaz, Joerg Doerr, Emitza Guzman, Mahmood Hosseini, Jordi Marco, Marc Oriol, Anna Perini, et al. The crowd in requirements engineering: The landscape and challenges. *IEEE software*, 34(2):44–52, 2017.

[14] John Grundy, Ingo Mueller, Anuradha Madugalla, Hourieh Khalajzadeh, Humphrey O Obie, Jennifer McIntosh, and Tanjila Kanij. Addressing the influence of end user human aspects on software engineering. In *International Conference on Evaluation of Novel Approaches to Software Engineering*, pages 241–264. Springer, 2021.

[15] John C Grundy. Impact of end user human aspects on software engineering. In *ENASE*, pages 9–20, 2021.

[16] Mariam Guizani, Igor Steinmacher, Jillian Emard, Abrar Fallatah, Margaret Burnett, and Anita Sarma. How to debug inclusivity bugs? a debugging process with information architecture. In *Proceedings of the 2022 ACM/IEEE 44th International Conference on Software Engineering: Software Engineering in Society*, pages 90–101, 2022.

[17] Emitza Guzman, Mohamed Ibrahim, and Martin Glinz. A little bird told me: Mining tweets for requirements and software evolution. In *2017 IEEE 25th International Requirements Engineering Conference (RE)*, pages 11–20. IEEE, 2017.

[18] Rashina Hoda. Socio-technical grounded theory for software engineering. *IEEE Transactions on Software Engineering*, 48(10):3808–3832, 2021.

[19] 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. IEEE, 2021.

[20] Hourieh Khalajzadeh, Mojtaba Shahin, Humphrey O Obie, Pragya Agrawal, and John Grundy. Supporting developers in addressing human-centric issues in mobile apps. *IEEE Transactions on Software Engineering*, 2022.

[21] Sanjay Krishnan, Jay Patel, Michael J Franklin, and Ken Goldberg. A methodology for learning, analyzing, and mitigating social influence bias in recommender systems. In *Proceedings of the 8th ACM Conference on Recommender systems*, pages 137–144, 2014.

[22] Jeff Larson, Surya Mattu, Lauren Kirchner, and Julia Angwin. How We Analyzed the COMPAS Recidivism Algorithm.

[23] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension. *arXiv preprint arXiv:1910.13461*, 2019.

[24] Ze Shi Li, Manish Sihag, Nowshin Nawar Arony, Joao Bezerra Junior, Thanh Phan, Neil Ernst, and Daniela Damian. Narratives: the unforeseen influencer of privacy concerns. In *2022 IEEE 30th International Requirements Engineering Conference (RE)*, pages 127–139. IEEE, 2022.

[25] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 2019.

[26] Walid Maalej and Hadeer Nabil. Bug report, feature request, or simply praise? on automatically classifying app reviews. In *2015 IEEE 23rd international requirements engineering conference (RE)*, pages 116–125. IEEE, 2015.

[27] Christopher Mendez, Lara Letaw, Margaret Burnett, Simone Stumpf, Anita Sarma, and Claudia Hilderbrand. From gendermag to inclusivemag: An inclusive design meta-method. In *2019 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, pages 97–106. IEEE, 2019.

[28] Danaë Metaxa-Kakavouli, Kelly Wang, James A Landay, and Jeff Hancock. Gender-inclusive design: Sense of belonging and bias in web interfaces. In *Proceedings of the 2018 CHI Conference on human factors in computing systems*, pages 1–6, 2018.

[29] Inês Nunes, Ana Moreira, and João Araujo. Conceptual modeling of gender-inclusive requirements. In *International Conference on Conceptual Modeling*, pages 395–409. Springer, 2021.

[30] Inês Nunes, Ana Moreira, and João Araujo. Gire: Gender-inclusive requirements engineering. *Data & Knowledge Engineering*, 143:102108, 2023.

[31] Arif Nurwidyantoro, Mojtaba Shahin, Michel Chaudron, Waqar Hussain, Harsha Perera, Rifat Ara Shams, and Jon Whittle. Integrating human values in software development using a human values dashboard. *Empirical Software Engineering*, 28(3):67, 2023.

[32] Lauren Olson, Emitzá Guzmán, and Florian Kunneman. Along the margins: Marginalized communities' ethical concerns about social platforms. *arXiv preprint arXiv:2304.08882*, 2023.

[33] Dennis Pagano and Bernd Bruegge. User involvement in software evolution practice: A case study. In *2013 35th International Conference on Software Engineering (ICSE)*, pages 953–962. IEEE, 2013.

[34] Sapna Parashar, Santosh Dhar, and Upinder Dhar. Perception of values: a study of future professionals. *Journal of Human Values*, 10(2):143–152, 2004.

[35] Matthew Pattison and Alex W Stedmon. Inclusive design and human factors: Designing mobile phones for older users. *PsychNology J.*, 4(3):267–284, 2006.

[36] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*, pages 28492–28518. PMLR, 2023.

[37] Margaret R. Roller and Paul J. Lavrakas. *Applied qualitative research design: A total quality framework approach.* Applied qualitative research design: A total quality framework approach. The Guilford Press, New York, NY, US, 2015. Pages: xviii, 398.

[38] Anthony Savidis and Constantine Stephanidis. Inclusive development: Software engineering requirements for universally accessible interactions. *Interacting with Computers*, 18(1):71–116, 2006.

[39] Shalom H Schwartz. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology*, volume 25, pages 1–65. Elsevier, 1992.

[40] Shalom H Schwartz. An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):11, 2012.

[41] Mojtaba Shahin, Mansooreh Zahedi, Hourieh Khalajzadeh, and Ali Rezaei Nasab. A study of gender discussions in mobile apps. *arXiv preprint arXiv:2303.09808*, 2023.

[42] Jack Shepherd. 23 essential twitter statistics you need to know in 2023. https://thesocialshepherd.com/blog/twitter-statistics#: :text=TwitterMay 2023.

[43] Ariana Tobin and Ava Kofman. Facebook Finally Agrees to Eliminate Tool That Enabled Discriminatory Advertising, June 2022.

[44] Virginia Tsintzou, Evaggelia Pitoura, and Panayiotis Tsaparas. Bias disparity in recommendation systems. *arXiv preprint arXiv:1811.01461*, 2018.

[45] Mihaela Vorvoreanu, Lingyi Zhang, Yun-Han Huang, Claudia Hilderbrand, Zoe Steine-Hanson, and Margaret Burnett. From gender biases to gender-inclusive design: An empirical investigation. In *Proceedings of the 2019 CHI Conference on human factors in computing systems*, pages 1–14, 2019.

[46] Watchful1. Subreddit comments/submissions 2005-06 to 2022-12.

[47] Grant Williams and Anas Mahmoud. Mining twitter feeds for software user requirements. In *2017 IEEE 25th International Requirements Engineering Conference (RE)*, pages 1–10. IEEE, 2017.

Appendix A

Details of the 50 apps

Table A.1: List of 10 business apps; Subreddit, Google play ID, and the Twitter Hashtag associated with each app.

App Name	**Subreddit**	**Google Play ID**	**Twitter Hashtag**
Adobe Acrobat	Adobe	com.adobe.reader	#adobe
Fiverr	Fiverr	com.fiverr.fiverr	#fiverr
Lightroom	Lightroom	com.adobe.lrmobile	#lightroom
LinkedIn	linkedin	com.linkedin.android	#linkedin
Microsoft Outlook	Outlook	com.microsoft.office.outlook	#outlook
Microsoft Teams	MicrosoftTeams	com.microsoft.teams	#microsoftteams
Trello	trello	com.trello	#trello
Webex	webex	com.cisco.webex.meetings	#webex
Workday	workday	com.workday.workdroidapp	#workday
Zoom	Zoom	us.zoom.videomeetings	#zoom

Table A.2: List of 10 entertainment apps; Subreddit, Google play ID, and the Twitter Hashtag associated with each app.

App Name	**Subreddit**	**Google Play ID**	**Twitter Hashtag**
Amazon Prime Video	AmazonPrimeVideo	com.amazon.avod.thirdpartyclient	#amazonprimevideo
Disney+	DisneyPlus	com.disney.disneyplus	#disneyplus
Hulu	Hulu	com.hulu.plus	#hulu
Netflix	netflix	com.netflix.mediaclient	#netflix
Paramount+	ParamountPlus	com.cbs.app	#paramountplus
Peacock TV	peacock	com.peacocktv.peacockandroid	#peacock
Pluto TV	Pluto_TV	tv.pluto.android	#pluto_tv
Roku	Roku	com.roku.remote	#roku
Tubi	TubiTV	com.tubitv	#tubitv
YouTube	youtube	com.google.android.youtube	#youtube

Table A.3: List of 10 financial apps; Subreddit, Google play ID, and the Twitter Hashtag associated with each app.

App Name	**Subreddit**	**Google Play ID**	**Twitter Hashtag**
Crypto.com	Crypto_com	co.mona.android	#crypto.com
Binance	binance	com.binance.dev	#binance
Google Pay	googlepay	com.google.android.apps.nbu.paisa.user	#googlepay
Paypal	paypal	com.paypal.android.p2pmobile	#paypal
CoinBase	CoinBase	com.remitly.androidapp	#coinbase
Robinhood	RobinHood	com.robinhood.android	#robinhood
Cash App	CashApp	com.squareup.cash	#cashapp

continues on next page

Table A.3 – continued from previous page

App Name	Subreddit	Google Play ID	Twitter Hashtag
Venmo	venmo	com.venmo	#venmo
Wealthsimple	Wealthsimple	com.wealthsimple.trade	#wealthsimple
Webull	Webull	org.dayup.stocks	#webull
Afterpay	Afterpay	com.afterpaymobile.us	#afterpay

Table A.4: List of 10 shopping apps; Subreddit, Google play ID, and the Twitter Hashtag associated with each app.

App Name	Subreddit	Google Play ID	Twitter Hashtag
Amazon Shopping	amazon	com.amazon.mShop.android.shopping	#amazon
Ebay	Ebay	com.ebay.mobile	#ebay
Temu	TEMU_Official	com.einnovation.temu	#temu
Etsy	Etsy	com.etsy.android	#etsy
Shopee	ShopeePH	com.shopee.ph	#shopee
Shopify	shopify	com.shopify.mobile	#shopify
Walmart	walmart	com.walmart.android	#walmart
SHEIN	SHEIN_	com.zzkko	#shein
craigslist	craigslist	org.craigslist.CraigslistMobile	#craigslist

Table A.5: List of 10 social media apps; Subreddit, Google play ID, and the Twitter Hashtag associated with each app.

App Name	**Subreddit**	**Google Play ID**	**Twitter Hashtag**
BeReal	bereal_app	com.bereal.ft	#bereal_app
Discord	discordapp	com.discord	#discord
Facebook	facebook	com.facebook.katana	#facebook
Messenger	facebookmessenger	com.facebook.orca	#facebookmessenger
Instagram	Instagram	com.instagram.android	#instagram
Snapchat	snapchat	com.snapchat.android	#snapchat
Twitter	Twitter	com.twitter.android	#twitter
WhatsApp	whatsapp	com.whatsapp	#whatsapp
TikTok	TikTok	com.zhiliaoapp.musically	#tiktok
Telegram	Telegram	org.telegram.messenger	#telegram

Appendix B

Examples of Inclusiveness Categories Vs Non-Inclusiveness

Table B.1: Examples of inclusiveness and non-inclusiveness from Google Play Store.

Category	**Quote**
non-inclusiveness	Downgraded from 3 stars to 2 stars! This app used to be great and they finally added the Sonic movie I so wanted to watch, but only a few minutes in and a few long minutes of ads, and to make things worst: The movie will rewind on it's own back to the beginning! It's unwatchable at this point! Glad to know I'm not the only one having this issue with the app! I'm considering just deleting this app and finding a better alternative!
other human values	If your VIEWS especially political and this covid scare then your account will be suspended for speaking your thoughts
demography	There getting like netfix if you dont live in states you lose alot of shows getting ready cancel it keeping taking away show I want to watch
fairness	Overall not bad but they suspended my account for no reason so un-susspend it and then we'll talk

continues on next page

Table B.1 – continued from previous page

Category	Quote
privacy	Too much personal data. This app takes way too many notes about you and your life. The feed is hyper specific and my advertisements are starting to correlate with my liked videos. I looked at permissions and this app requires access to call logs, message logs, phone ID and security access. It also has access to saved payment methods. This app is malicious and I don't know what the CCP wants with that data access, but I'm not okay with Commies spying on my life. Screw bite dance
technology	Tiktok is good app but sometimes we can't open or login our I'd in another device
usability	Annoying text.. button follow user when watching video.. please remove text on any video.. cannot satisfy video watching nowadays.. too many text crowded on video ...

Table B.2: Examples of inclusiveness and non-inclusiveness from Reddit

Category	Quote
non-inclusiveness	Debit Card on CA If I used my linked debit card on my cash app but never added money to my account from it. Instead used the Cash Card for my purchases, would that work? Would that appear on my bank statements as Adding Cash to my account or would it just say what things I used it for. Am I wrong to assume that if my cash account were for some reason to hit zero it would pull funds from my linked card without warning?

continues on next page

Table B.2 – continued from previous page

Category	Quote
other human values	NETFLIX is losing subscribers! I keep reading articles about NETFLIX losing subscribers. IMO the best thing about NETFLIX are seriouses and documentaries. Obviously, TV shows are polarizing in the sense people will either like it or not. However, documentaries are supposed to be informative and thought provoking. Unfortunately, most of the controversial documentaries are ridiculously predictable. All you have to do is read the title and you already know what slant it will take, especially if there is a political aspect. I'm sure documentaries exist that take a libertarian point of view, or simply something that doesn't align with the Democrat party. Or is all diversity in thought dead? Why can't I watch something that challenges my point of view? Documentaries, are a huge part of my NETFLIX experience, but without new fresh point of views, I may have to leave. Me leaving would have nothing to do with HBO+, Disney+ or any other streaming service.
demography	Problem with subscribing to discord nitro, Help? I used to be a nitro subscriber and I had to upload my billing address, it wont recognize my location even though Ive tried everything, I wish it would say something like address invalid or incorrect postal code would help a lot
fairness	my account got suspended I got a mail 2 days ago saying that my account got suspended for human exploitation (which i didn't do) and even banned my alt which was just dead and rarely used it. I had a request review mail but clicking it just led to Facebook homepage. I went to Instagram suspended forms to fill it out but when i type my id there it says id is not valid. I literally don't know what to do now. Wait for 30 days? Will i get my account back? My account now to everyone else pov is just Instagram user. I really want my account back. All my photos and memories are stored there.
privacy	Is CoinBase worth it? I know its the go to platform for beginner crypto purchasing, but Im not comfortable with the lack of service/complaints from everywhere Ive heard and already experienced. I verified my bank on CoinBase Pro the way they told me to just for them to give me a suspicious activity warning. Im annoyed and confused, any help appreciated.

continues on next page

Table B.2 – continued from previous page

Category	Quote
technology	Please revert the update. The new update made everything extremely sluggish on android. So many neat little details are gone (like the chat easing up as new messages are sent, and emote names being highlighted in blue). It doesn't look much better and the previous update was so much more pleasant to look at. Please give us the option to either revert the update or change the layout back to what it used to be.
usability	Notification noise from *inside* app on Android is an issue. There's nothing worse than being inside telegram and hearing that stupid plonk sound from someone else messaging me (or even worse, when I'm not in any particular chat and just looking at the list). I can't find anything about this noise or how to get rid of it but I really need to. Thanks.

Table B.3: Examples of inclusiveness and non-inclusiveness from Twitter

Category	Quote
non-inclusiveness	Original Lenovo Lp3 pro TWS Earphones Wireless Bluetooth 5.2 Earbuds HIFI Sound Noise Reduction HD Call Low Latency Headphones
other human values	You officially lost a customer you are a big time loser @facebook learn to respect freedom of speech! #socialmedia #marketing #socialmediamarketing #digitalmarketing #instagram #branding #business #twitter https://t.co/nakhNoHEgm
demography	I feel like a senior citizen trying to figure out this BeReal app
fairness	@instagram why is your recovery process terrible? My husband's account was hacked over a year ago and still nothing has been done. It is still up?!! #Instagram
privacy	Hey @bobby_tetsch they are tracking you on #Facebook and it will not stop and might even get worse if you dont speak up ! Also you quoted Acquisiton news by end of December ??? WAS THAT A LIE TO BAG ALL OF US AS WELL ? JUST SAYING https://t.co/5G0874kKXU

continues on next page

Table B.3 – continued from previous page

Category	Quote
technology	Anyone else having trouble using #Twitter on firefox? My FF is updated, everything else is running smoothly on there, but my twitter is not loading anything... I just downloaded google chrome to try it on here, and here it works... I don't want to use GC though...
usability	Who has time to use a mouse or trackpad when responding to emails? Use this great shortcut to reply to your Outlook emails without your fingers needing to leave the keyboard https://t.co/AZ4IFFiPkv #Microsoft #Outlook #OullookShortcuts https://t.co/AZ4IFFiPkv

Appendix C

Memos

Fairness:
From the data, it can be inferred that many software companies enforce rules and policies that often lead to discrimination against certain users, denying them access to the app. Such behavior leaves numerous users feeling excluded from using the app. There is also a lack of effective customer service offered by these software companies. Users express frustration, feeling that the support service is either unhelpful or fails to address their needs. This lack of proper customer support is perceived as unfair and creates a sense of exclusion from the app.

Terms and Conditions:
Based on the data, there is an increasing trend in companies to update their terms and conditions, which directly impact users' ability to use the software. Many users report experiencing bans for possible violations of these updated regulations, which they were either unaware of or did not perceive as a violation. These bans are the result of decision-making algorithms that are incorporated in the updated terms and conditions. This raises concerns among users regarding the perceived fairness of these terms and conditions, and they feel excluded due the perceived unfair behaviour.

Recommendation: The data indicates that users receive content or features that are not aligned with their personal preferences. They constantly update and refine their settings, but the app does not respond accordingly. This discrepancy creates a perception of exclusion, where users feel that the app does not cater to their individual needs. Key concerns are lack of personalization and frustration with unexpected ads. Thus, users feel like the app is not being inclusive to them.

Service:
A substantial number of users express concerns about their interactions with customer service. Users report facing difficulties in accessing the app or specific features. Often, users encounter bans and require a review process, but the automated nature of these reviews leads to a perception of unfairness. Users find this lack of personalized support to be unjust. They strongly express a more interactive and human-centric support system where they can provide context and receive personalized assistance for any issue. The lack of such a support system creates a sense of exclusion, particularly in terms of fairness.

Technology:
The data highlights users complaining about their ability to access a particular app due to technological constraints. Many users express frustration at being excluded from using the app due to a mandatory reliance on another technology. It is apparent that there is a lack of awareness that not all users may have access to or wish to use the connecting technology. The assumption that all users possess this technology creates a direct impact on their ability to use the app, leading to a perceived lack of inclusiveness.

Website:
The data presents concerns revolving around the authentication process within the app. Specifically, users report an inability to log in without an account on another website. Their attempts to log in lead them to authentication prompts, a service they either do not have or choose not to use. This authentication requirement poses a significant barrier, effectively excluding users who do not use this specific website. This requirement appears to be enforced by the development team. Such enforcement generates a feeling of exclusion among users.

Network:
"the app never works with my wifi" Many users report an inability to access the app while connected to a particular network, while it functions seamlessly on other networks. This discrepancy creates a direct sense of exclusion, as users find themselves unable to use the app in their preferred setting. Furthermore, some users relate this issue to recent updates that may not support their older network.

Device:
The data shows a recurring concern among users related to device requirements. Many users report instances where the app either fails to function on their devices altogether or abruptly stops working. These issues stem from various factors, including companies discontinuing support for specific devices, updates that make the app incompatible, or a lack of consideration for particular devices during development. As such. users experience a sense of exclusion from the app.

Privacy:
The data analysis indicates a number of privacy concerns from users. Many software apps require providing highly sensitive information, including bank account pins, government IDs, and credit card details, in order to access certain features. This demand for personal data creates a sense of privacy breach among users, leaving them feeling frustrated and excluded. The users express concerns about the safety of their information and experience a direct impact on their ability to use the app.

Security:
The data presents issues concerning security incidents reported by users. Hacking incidents are common, leading to users facing difficulties in account recovery or access. This not only instills a sense of fear and insecurity but also creates a feeling of exclusion, as users doubt the app's ability to protect their accounts. Users also encounter unusual activities, and the potential threat of scams further increases these concerns.

Demography:

The data presents a diverse range of concerns directly linked to demographic characteristics. These demographic factors significantly impact the user experience, leading to instances where users feel excluded or underserved.

Location:
Many users report either being denied access or experiencing significant difficulties in using the app due to their specific region. This geographic limitation significantly impacts the user experience, as they find themselves unable to access certain features or, in some cases, the entire app. This generates a feeling of being excluded from the app.

Age:
The data uncovers a series of age-related concerns among users. These concerns highlight users who initially created accounts with inaccurate birth dates and now fear a potential ban. Newly imposed age verification rules are developing a fear of losing their account. Additionally, there are complaints regarding apps that appear to have an abundance of features that do not cater to older demographics. These age related concerns collectively underscore a sense of exclusion and a need for greater inclusiveness in the app.

Gender:
The data suggests apps focusing on a particular gender. Specific features in these apps appear to be more preferred by one gender over others. This discrepancy has led individuals of other genders to feel excluded, as they perceive a lack of representation and relevance in the app.

Language:
The data analysis highlights many language related concerns. Users frequently describe disappointment when they encounter difficulties in locating their specific language settings in the app. The absence of their preferred language, while other language options are available, contributes to a sense of exclusion and frustration.

Socio-economic status:
The data infers frustration and a sense of exclusion among users when encountering apps that offer premium features or require payment to access basic account functionalities. These economic barriers create a significant challenge for users and contribute to a sense of exclusion while using the app.

Usability:
Users express frustration when the usability of the app hinders their ability to have a seamless and enjoyable experience. The primary concerns revolve around visual and audio-related problems, which appear to be the primary concern of the users.

Visual:
Users consistently express a preference for specific interfaces and color schemes but frequently encounter challenges in finding these options in the app. This discrepancy between user preference and available customization options significantly affects their experience with the app. Additionally,

many apps appear to lack accessibility features, which further creates difficulties for users and fosters a sense of exclusion. Key concerns: difficulty in finding options, customization, and accessibility.

Audio:
Many users experience issues with the app's audio functionality. Users frequently report various sound related issues, most notably an inability to hear anything from the other side. They also encounter problems with low audio volume or complete inaudibility. These issues appear to stem from updates or device/microphone compatibility problems. Sound-related issues often originate from recent updates or potential compatibility issues between the app and users' devices or microphones.

Other human values:
A recurring theme has emerged regarding users' perceptions regarding other human values. It is evident that a significant number of users feel that the app tends to favor a specific group of individuals who hold particular beliefs and restrict the others. These include restriction for freedom of speech, social justice issues, and welfare of friends and family which can be denoted as benevolence. There is a perception of favoritism and diverse belief systems that emerged from the data. At the same time, there is a lack of collectivist features which users perceive as exclusion.

Freedom:
Under human value related concerns, a prominent issue surrounded users expressing political or social views within the app. A significant number of users report instances of being banned or having their content restricted when sharing political opinions. This pattern is leading users to feel unwelcome and marginalized within the app community. Key concerns include censorship of political views and restriction of freedom of expression. Such issues are making users feel like they are unwelcome in the app.

Social Justice:
Users express a lack of inclusion related to their background and ethnicity. It is evident that many users repeatedly feel their religious or ethnic background is restricting them from fully experiencing the apps features. They feel the app does not let them use the features due to bias against them. Thus, this violates the social justice aspect and users feel excluded from the app.

Benevolence:
The data indicates that users from different cultural backgrounds anticipate collective features within the app, such as the ability to add family members. At the same time, they are anxious about the lack of child safe features and content. The users are frustrated as they are worried about the welfare of their near and dear ones. These gaps result in the users' feeling excluded.