

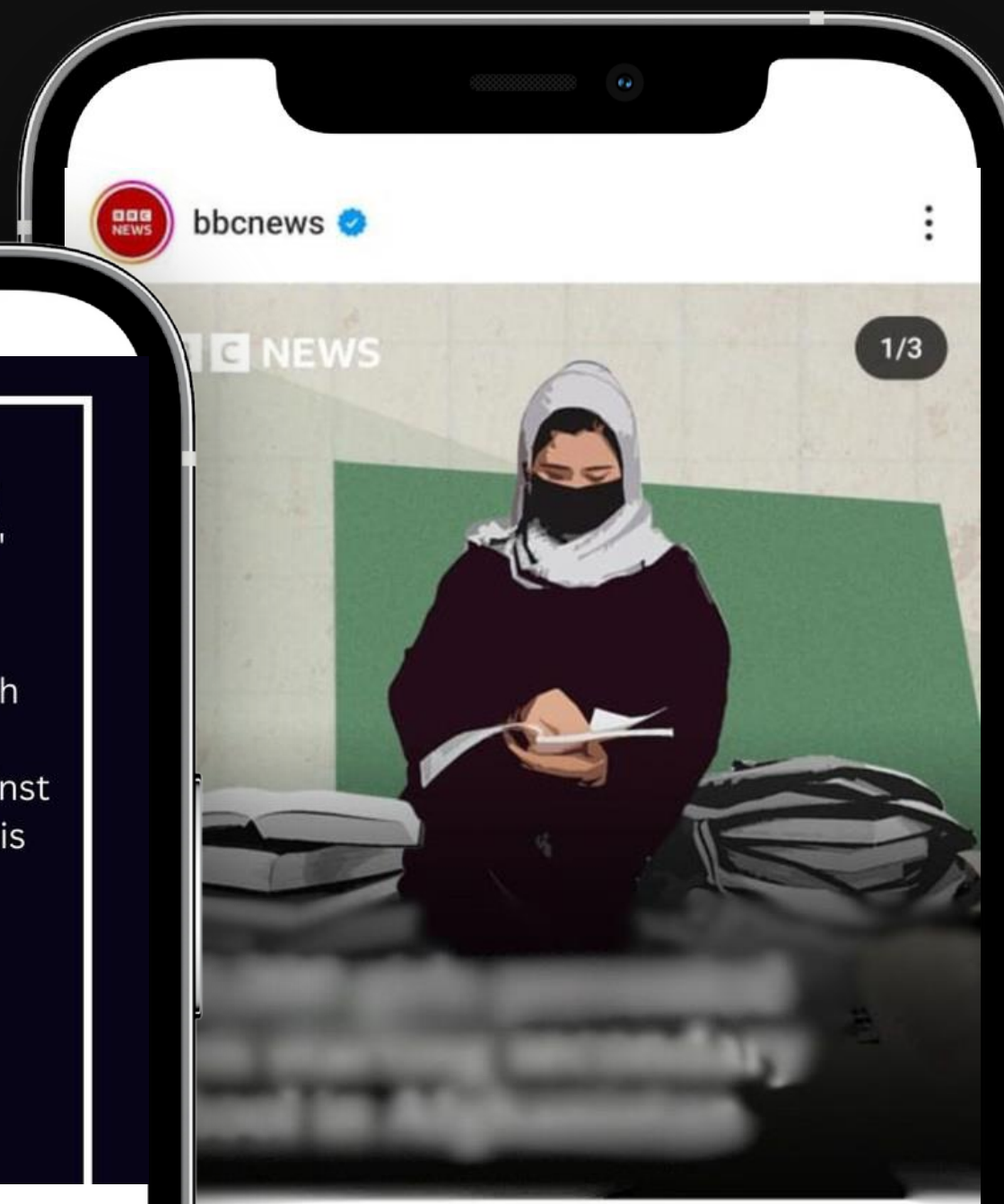
# SafeSpeak

## Detected Hate Speech:

"Those [ethnic group] people are nothing but criminals and leeches on our society."

## Explanation:

This statement was flagged as hate speech because it uses offensive language and negative stereotypes to discriminate against an entire ethnic group. Statements like this promote hatred and intolerance.



# WHY WE NEED IT ?

A young woman of color, vocal about issues of racial justice online, faces a barrage of hateful messages on social platforms.

The attacks range from insults about her appearance to threats against her safety. Reporting tools on social media platforms feel ineffective. The constant negativity takes a toll on her mental health, making her question her online presence and activism.





## What?

We are developing a plugin that filters hate speech and offensive content from social media platforms.



## Why?

Hate speech and offensive content can create a negative and harmful online environment.



# 36%

of online users have been targeted by hate speech based on their religion.

---

# 78%

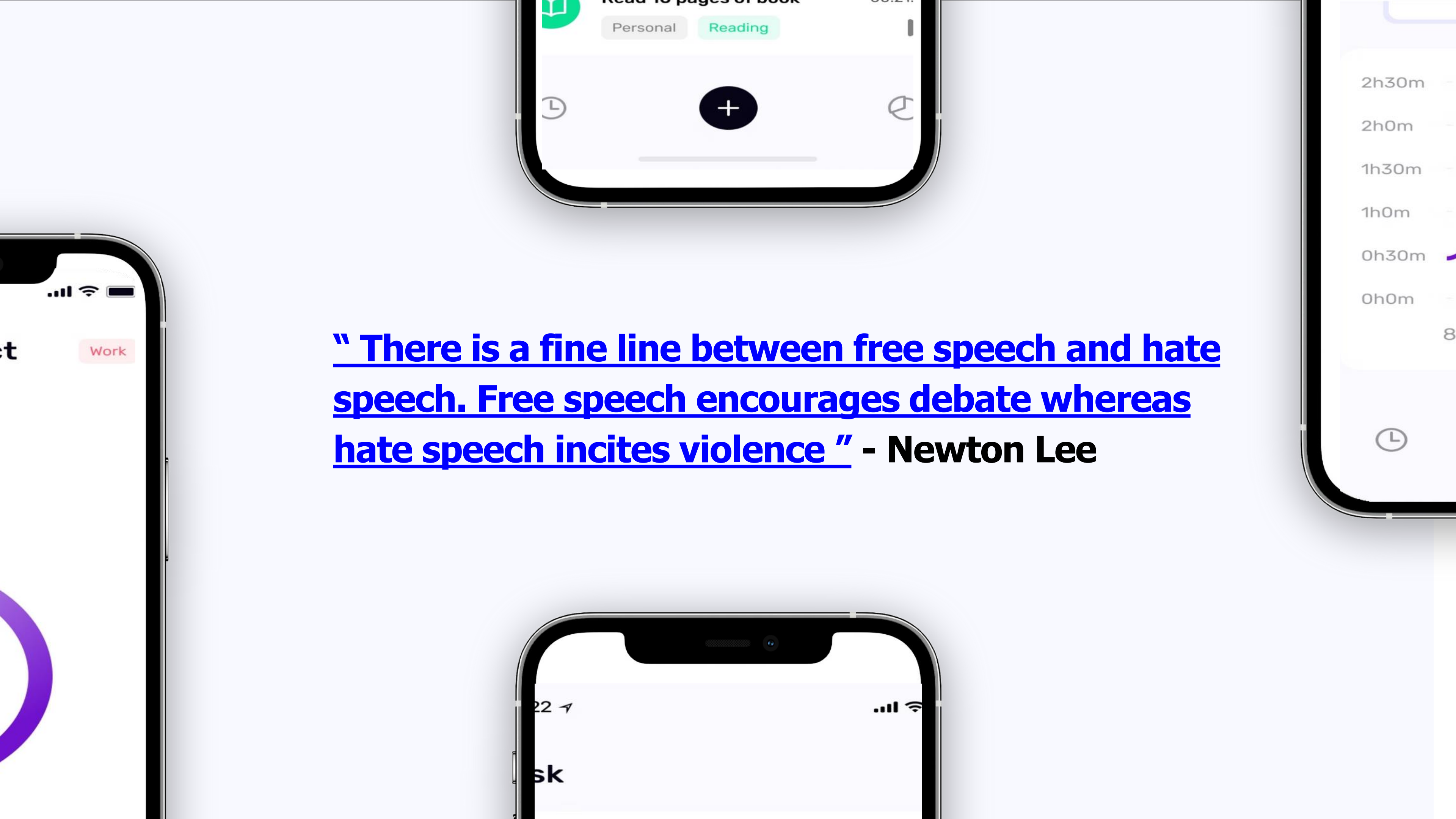
of internet users globally believe social media platforms should be more proactive in removing hate speech.

---

# 67%

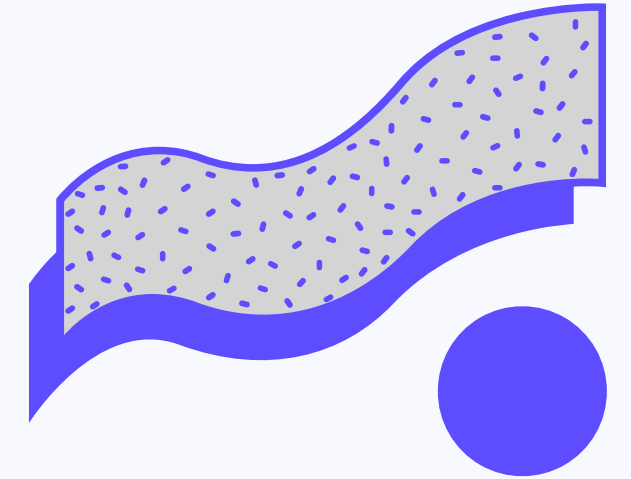
of social media users globally believe they should have more control over the content they see online.





**“ There is a fine line between free speech and hate speech. Free speech encourages debate whereas hate speech incites violence ” - Newton Lee**

# Market Need



Rising concerns about online hate speech and harassment.

Increased awareness of the negative impact of online negativity on mental health and social dynamics.

Growing demand for content moderation tools by social media platforms.

# SIZE OF MARKET

1.

TA  
M

2. SAM

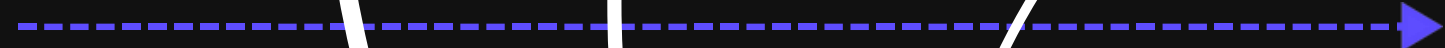
3. SOM



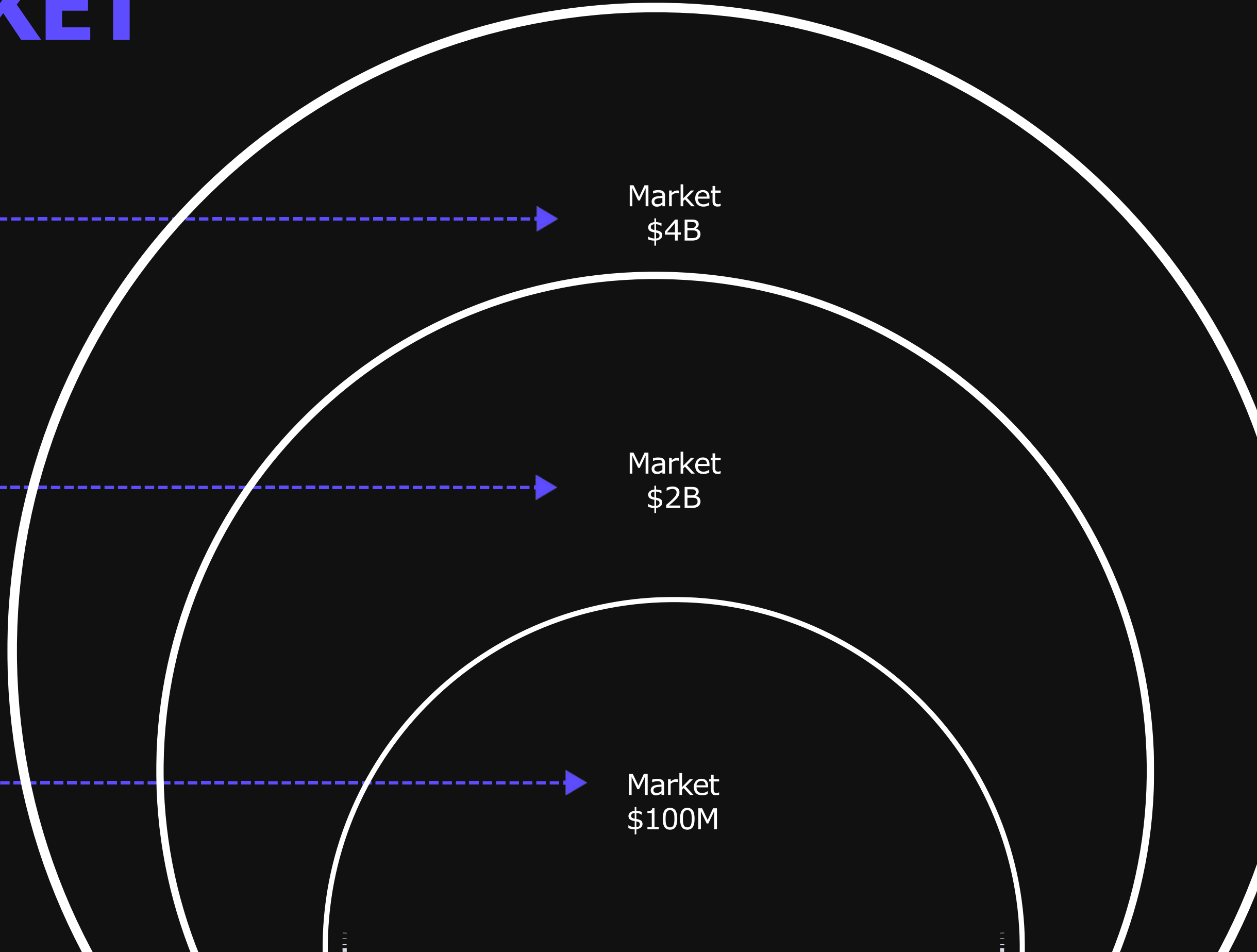
Market  
\$4B



Market  
\$2B



Market  
\$100M



# Target Audience



**Social  
Media Users**

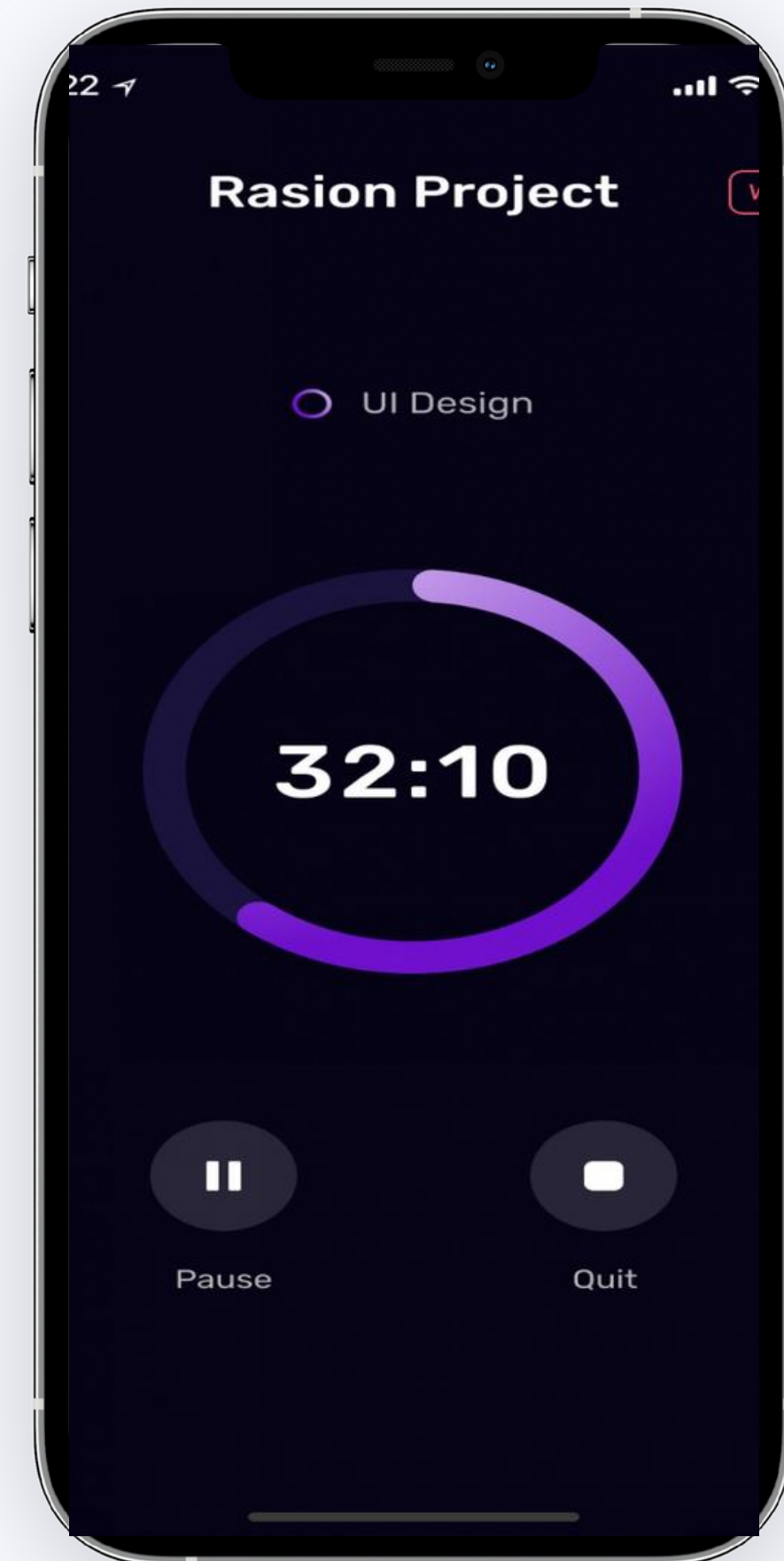
**Parents  
and  
Guardians**

**Content  
moderators**



# Business model

Our hate speech recognition plugin operates on a subscription-based model, providing a reliable and scalable solution for businesses and online platforms to combat hate speech and maintain a safe online environment. With our plugin, users can easily integrate hate speech detection capabilities into their existing systems and applications



# Competitive advantages

1

Instead of just flagging content as hate speech, your tool could provide explanations and insights into why certain text was classified as such.

2

Different communities and platforms may have varying definitions of what constitutes hate speech.

3

Instead of just flagging past content, your tool could provide real-time monitoring and alerts for hate speech as it occurs, enabling faster response times.

# Product features

**Explainable AI**



**High Accuracy**



**Customizable Definitions**



**Integration**



**Multilingual Support**



**User Feedback Loop**



**Real-time Detection**



# TEAM MATES



**Gaurav**  
Team Leader



**Gurwinder Singh**  
Programmer



**Deepak**  
Programmer



**Jatin**  
Business Area

Company Name

SafeSpeak

Contact

[deepak747@gmail.com](mailto:deepak747@gmail.com)

# Thank You

