



# Community Chat Takedowns

An end-to-end design solution for a 0 to 1 product  
in Facebook Messenger

JULY 2022 | CONFIDENTIAL



**Background**

# AGENDA

1

## Background

How long was this project?  
What is Community Chats?

2

## Problem

What do we know about the problem?  
What problem were we solving?

3

## Approach & Solution

How did we approach the solution?  
What considerations and tradeoffs did we make?

5

## Results & Learnings

What came out of this project?  
What would I do differently?

# PROJECT TIMELINE

Time frame to understand the problem, explore designs, deliver feasible solutions, build, and launch MVP.

## MONTH 1



### Onboard & Understand

Meet the new team, onboard to the new product, understand the users and problem space.

## MONTH 2



### Ideate, Diverge, Converge

Explore ideas for ways to solve the problem, meet with overlapping teams to understand considerations, share ideas with team and receive feedback, iterate.

## MONTH 3



### Propose & Review

Propose recommended solution to cross-functional and leadership teams, understand any gaps or improvements needed.

# PROJECT TIMELINE

Time frame to understand the problem, explore designs, deliver feasible solutions, build, and launch MVP.

## MONTH 4



### Iterate and Build

Iterate on feedback received from cross-functional partners and leadership team, propose modified solution, hand off and collaborate with engineers on the build.

## MONTH 5



### Build and Iterate

Close collaboration with Engineering and Product for technical constraints and support during build.

## MONTH 6



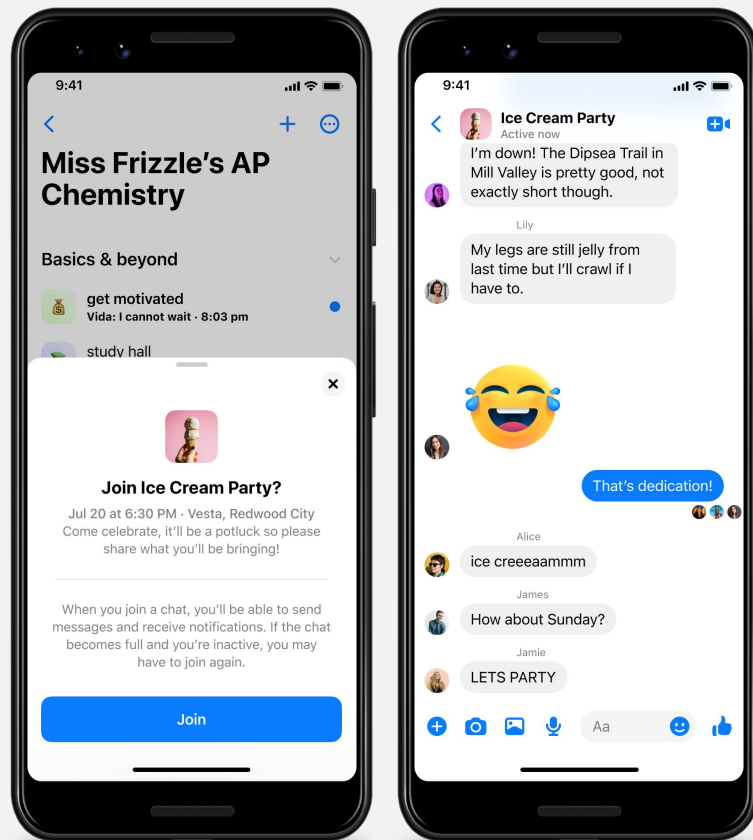
### Launch

Support any edge cases during build, test, bug bash, deliver and launch MVP of the new Community Chats product.



## COMMUNITY CHATS

Community Chats is the latest Messenger product that enables communities to connect via chat, audio and video with a meaningful audience: the communities that matter to them.

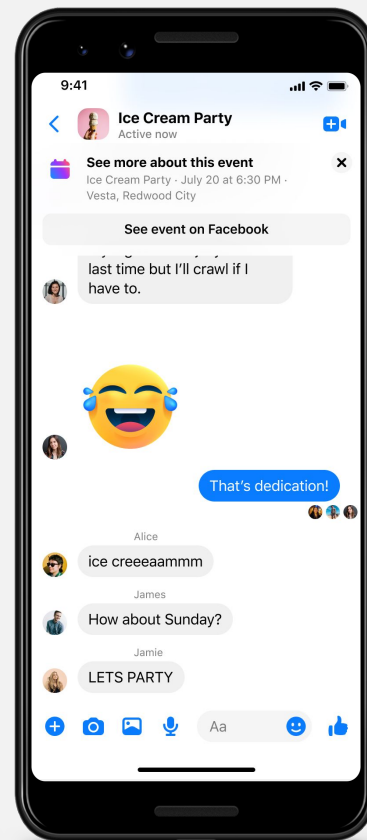
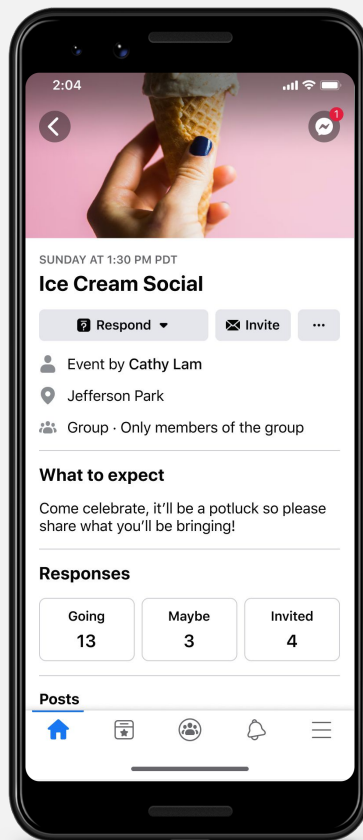




## ENFORCEMENT IN COMMUNITY CHATS

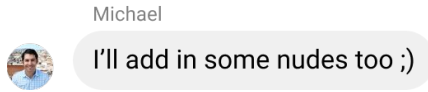
Community Chats integrate with Facebook Groups, where chat threads are analogous to posts in a group and messages are like comments. For this reason, Community chats use the open feed enforcement model and are considered a public surface.

Content in that violates Community Standards must be enforced upon and taken down.



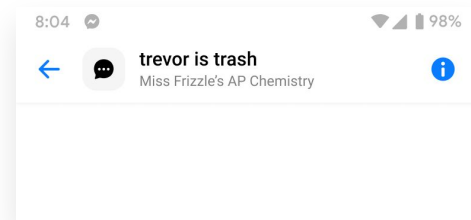
## Content that can be considered violating in Community Chats are messages and chats

---



### Messages

Chats can contain messages that are violating, be it text or images.



### Chats

A whole chat is considered violating if any group metadata is violating.



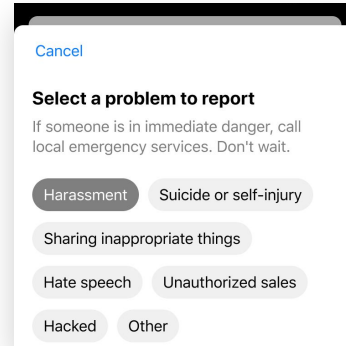
## Violations can be detected or reported

---



### Detection

Violations can be detected via MMS, a scanning service that matches illegal images against a bank of images.

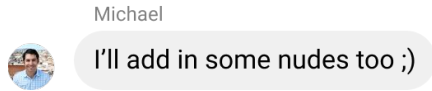


### Reports

Violations can also be reported by members, which flags content to be reviewed by Operations representatives.

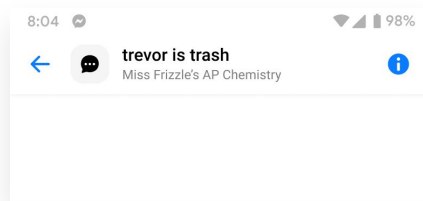
# Violations can result in...

---



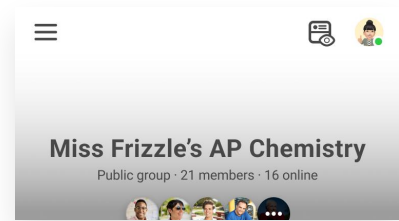
## Message takedowns

If a message (text, media) is found to be violating either through detection or through human review, it can be removed.



## Chat takedowns

If a chat contains messages that are violating, be it text or images, it can be removed.



## Group Takedowns

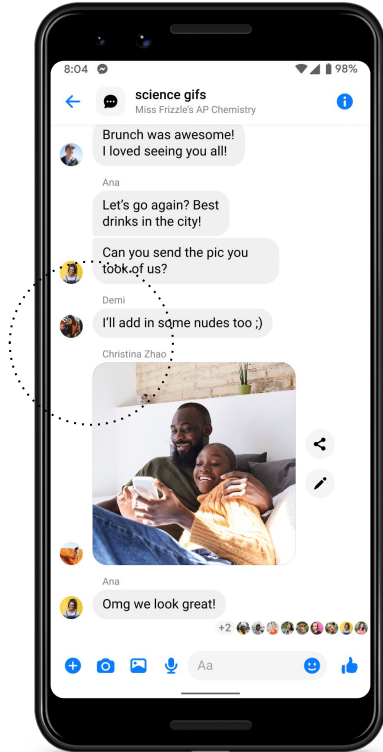
If the violator was an admin, the group gets a Group strike. If a Group passes the group strike threshold, the entire group can be taken down.

## PROBLEM

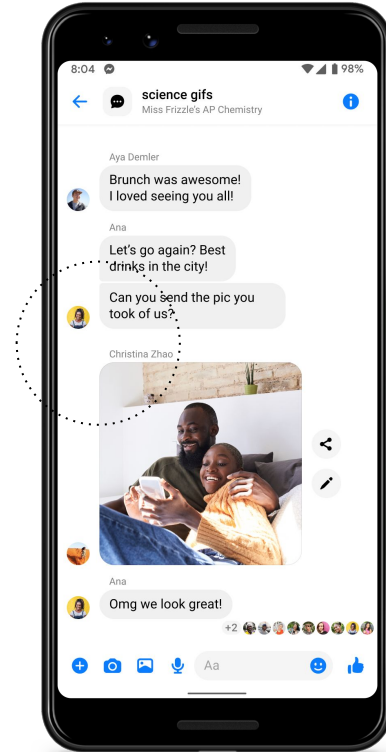
How might we help users  
**understand that content has been  
removed due to a violation  
of Community Standards?**

# Today: Message Takedowns

Today when a message is found to be violating, it is simply removed with no context.



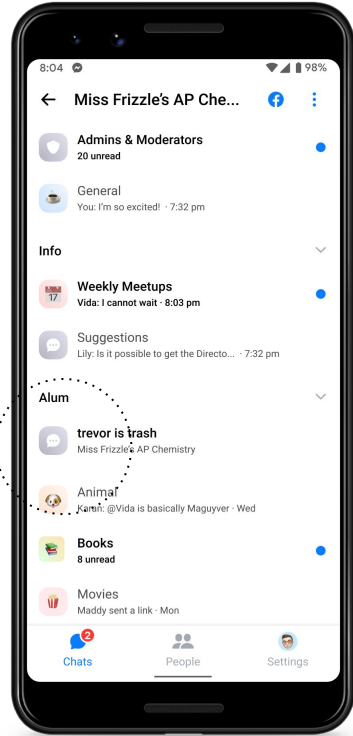
Violating message detected



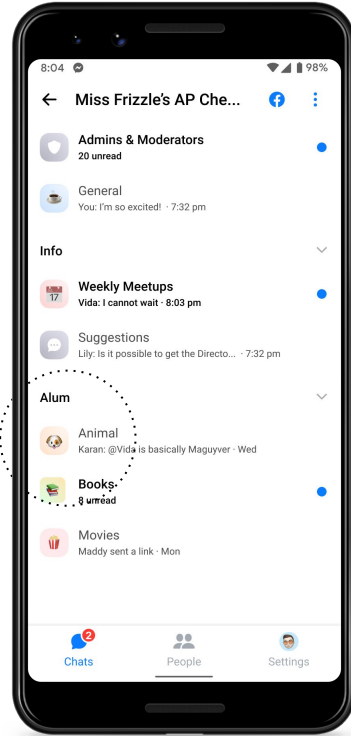
Post-action: message deleted

# Today: Chat Takedowns

Today when a chat is found to be violating, it is simply removed with no context.



Violating chat detected



Post-action: chat deleted

# Approach

# Goals



## Inform users that a message or chat was removed

Help inform thread participants that a message was previously there to help provide contextual continuity.



## Norm Community Standards

Educate both actors and bystanders that we have a set of Community Standards and that we take action when these are violated.

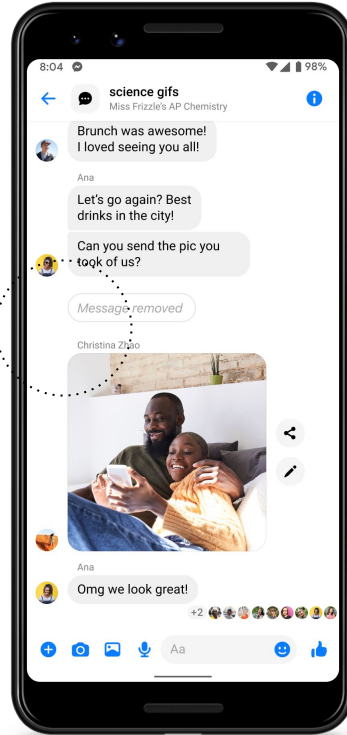


## Educate respectfully

Educate thread participants without shaming the actor or creating pile-ons. In addition, inform and differentiate content removal from Meta vs. from admins.

# Explorations

Early iterations for message takedown treatment

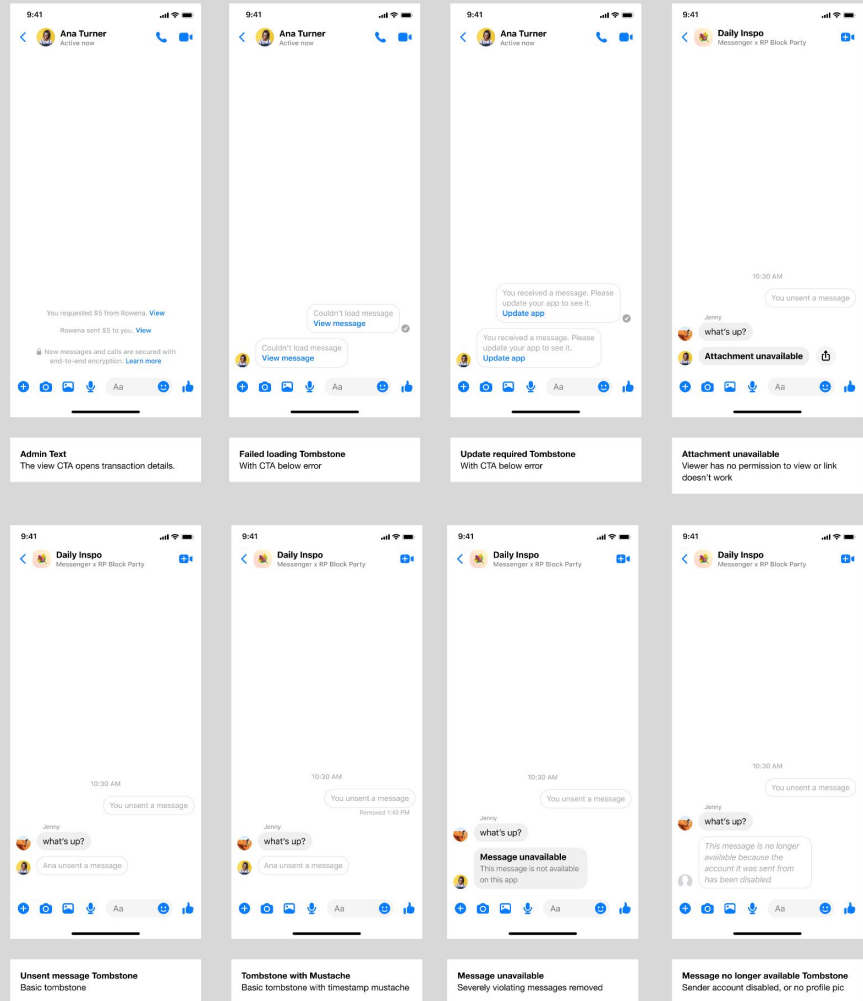




# CONSIDERATIONS

We performed an audit of existing tombstone patterns as part of the process in deciding the visual treatment for removed messages. The end solution builds on top of existing patterns in the app and also optimizes for visual consistency.

## EXISTING TOMBSTONE / ADMIN TREATMENTS TODAY



**Solution**

# Proposed: Message Takedowns

A violating message will not only be removed, but will also be tombstoned



Violating message detected



Post-action: message tombstoned



# Message Tombstone



## Tombstone treatment

A tombstone not only reuses an existing pattern from Facebook, but also helps maintain some continuity of the thread.

Furthermore, replacing a deleted message with a tombstone helps reinforce norms of Facebook's Community Standards.



## Anonymized profile

The image of the violator's profile would be converted into the default profile image, protecting the privacy of the violator and prevent shaming.

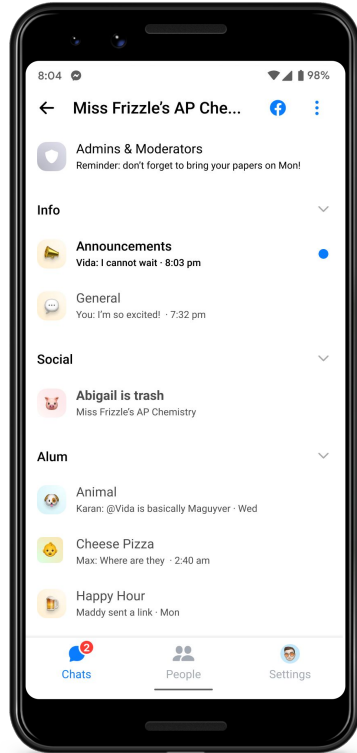


## Voice standards

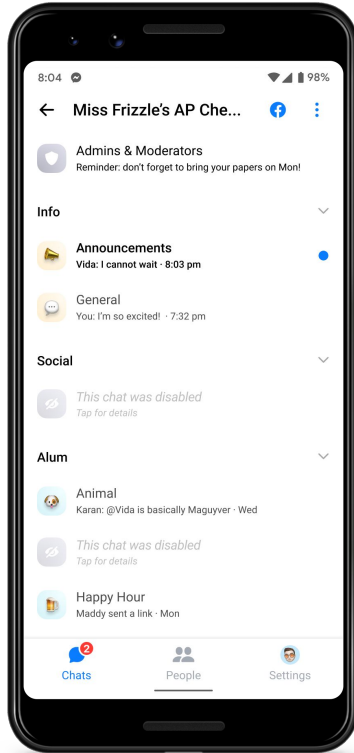
When the user is taking action - the subject of the sentence is the user. When Meta is taking action - the subject of the sentence is the content.

# Proposed: Chat Takedowns

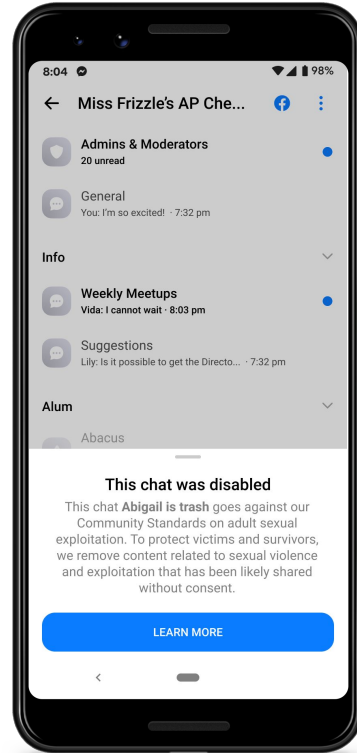
A violating chat will not only be removed, but will also be tombstoned



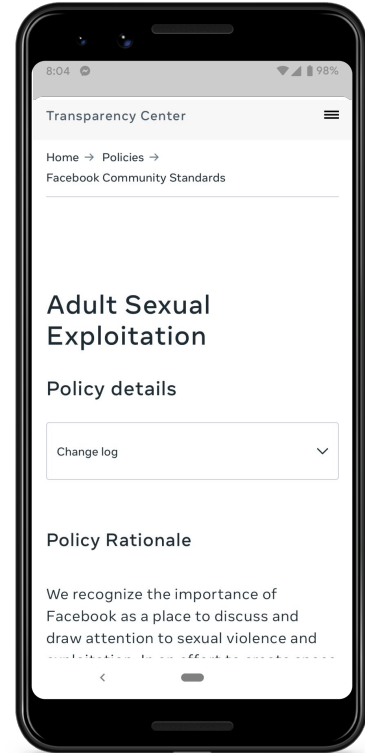
Violating chat detected



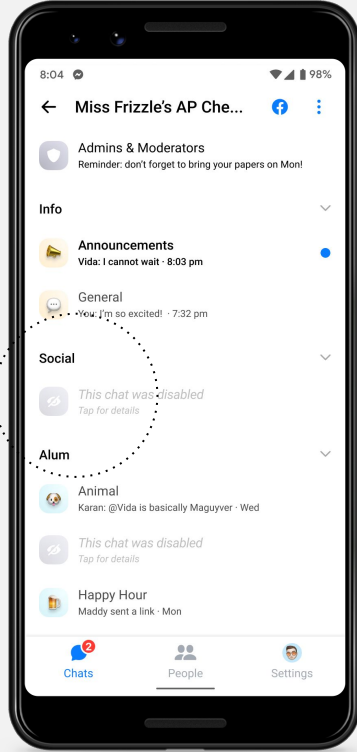
Post-action: chat tombstoned



Policy snippet on tap for details



Learn more in Transparency Center



# Chat Tombstone



## Tombstone treatment

Leaving a trace of the chat in the list promotes norms and helps users maintain context of a thread they may have been active in.

A removed chat will be replaced with “This chat was disabled” and will become inactive. The ability to learn more would be accessible via “tap for details”.



## Tombstoned content

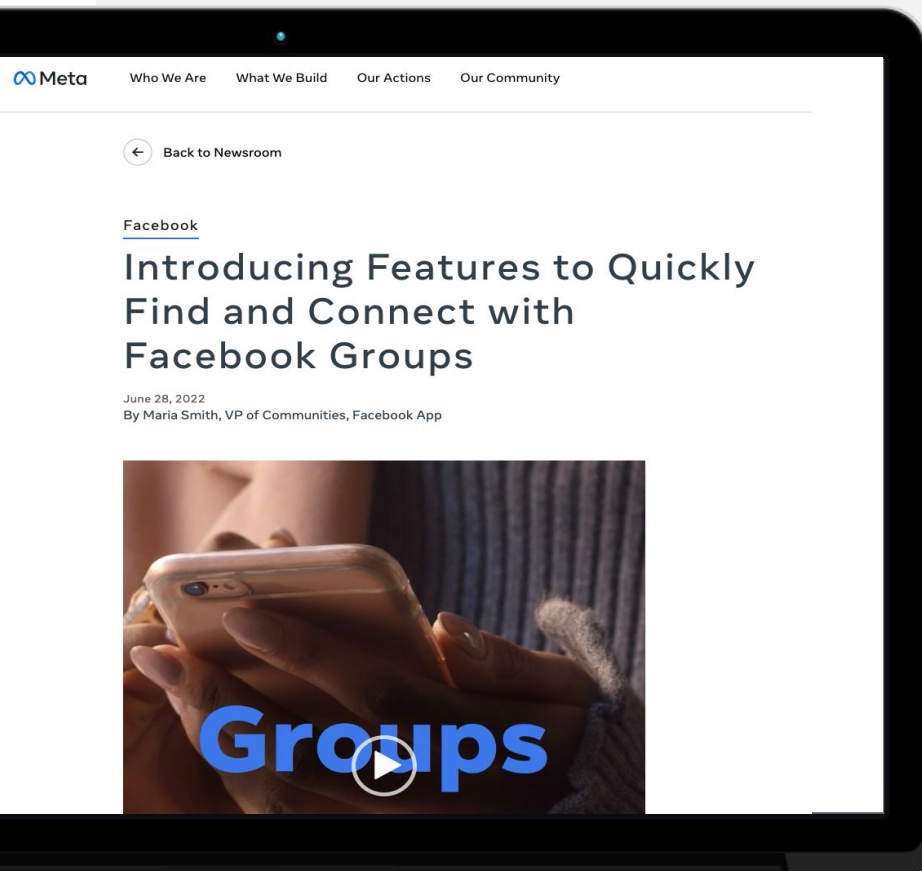
In order to future-proof this solution, we had to consider how to inform users more about why their chat was disabled in a scalable way for when the chat preview line disappears.



## Policy details

Tapping on a disabled chat would bring up a dialog providing more detail about the violation using policy snippet language, helping norm Community Standards.

# Results & Learnings



# MVP Launch & Takeaways

The team worked tirelessly to launch the MVP for enforcement experiences in Community Chats launched in June 2022. This new product is a focus and investment for the company to increase engagement within Facebook.

## Proactive communication with overlapping teams

Because Community Chats crossed multiple products and teams, this project was complex and challenging, requiring me to be tuned into all overlapping teams' future vision and upcoming changes in order to anticipate the long-term scalability of my design solutions.

## Agile collaboration while building

Launching a 0 to 1 product meant there was no blueprint or precedent to follow. As the lead designer on this project, it was necessary to be agile and quick to adjust design solutions as new edge cases and Eng constraints arose.



**THANK YOU!**