

Indicate what categories of sensitive content you want to block from tasks

- hate speech
- cyberbullying and harassment
- discrimination
- abuse and profanity
- misinformation and deceptive content
- illegal activities and promotion
- other

click to fill in the blank