

How to Moderate Usernames Like a Pro

“Hello, my name is...”

We’ve all worn them, whether at a conference, meet and greet, or first day at a new job: those colorful “Hello, my name is...” stickers are everywhere. They serve a valuable purpose, despite how silly they look. Our names are core to our identities, and the first word we use when asked to describe ourselves. We form connections to people when we use their names. One of the most powerful things a person can do is use your name when they speak to you.



And we don’t even get to choose our real names! So imagine the power inherent in selecting a username in a game, social network, or app. We select account names based on our personalities, our passions, our sense of humor. We choose names because they resonate with us, because they project a certain image, because they reflect our particular subculture.

Users don’t take their display names lightly – and community managers and moderators shouldn’t either.

Setting the tone



Now, imagine a product full of screen names like “TehKiller,” “Screw0ff,” and “URcr0tch.” Not very inviting, are they? For a new user, display names can set the tone for their entire experience in your product. And if that experience doesn’t welcome them, there is a

good chance they won't return. Today's consumer has endless options and limited time – so it's crucial that you grab their attention right away, and keep it.

Consider this: creating a username is also the first impression users have of your site security. Users who are inclined to create high-risk names are also inclined to post high-risk or abusive User Generated Content (UGC). So it's key that you set community expectations and enforce community guidelines early in the experience.

An effective username moderation strategy can help you set your product's tone from the first time a user logs on.

Getting started

Before you can design a username workflow, consider how you handle UGC in your community.

Usernames aren't easy to filter

First, keep in mind that username moderation is technically different from text moderation. As we wrote in [What is the difference between a profanity filter and a content moderation tool?](#), don't assume that a simple profanity filter is equipped to detect and flag offensive language in usernames. Unlike other UGC like chat, messages, comments, and forum posts, usernames rarely consist of "natural" language. Instead, they're made up of long strings of letters and numbers – "unnatural" language.

Most profanity filters lack the complex technology to filter usernames accurately, but some moderation tools are designed to adapt to all kinds of different content.

So even if you're using an [internal solution](#) for text or even image moderation, consider a content moderation system for username moderation. This will ensure accuracy and maximize community/brand protection.

Label content by risk and topic

Another reason your community and brand will benefit from using more than a profanity filter. If you divide your content based on risk levels and specific topics, you can easily sort usernames into buckets, allowing you to automatically approve, reject, or manually review them in a timely and efficient manner.

On a simple level, imagine three basic risk buckets: low, medium, and high. Then, based on your Code of Conduct, decide which topics you want to filter (like hate speech, profanity, PII, in-game assets, etc).

Every community is different and will have different tolerance levels. For instance, battle games may allow names that include violent imagery.



Don't forget to display your Code of Conduct during the account creation process!

COPPA applies to usernames, too

If your product is targeted at under-13 users in the US, you'll need to abide by the Children's Online Privacy Protection Act (COPPA). There are several layers to COPPA compliance (check out our [beginner's guide](#)), but in the moderation space it's critical that you filter personally identifiable information (PII) like first and last names, email addresses, and phone numbers.

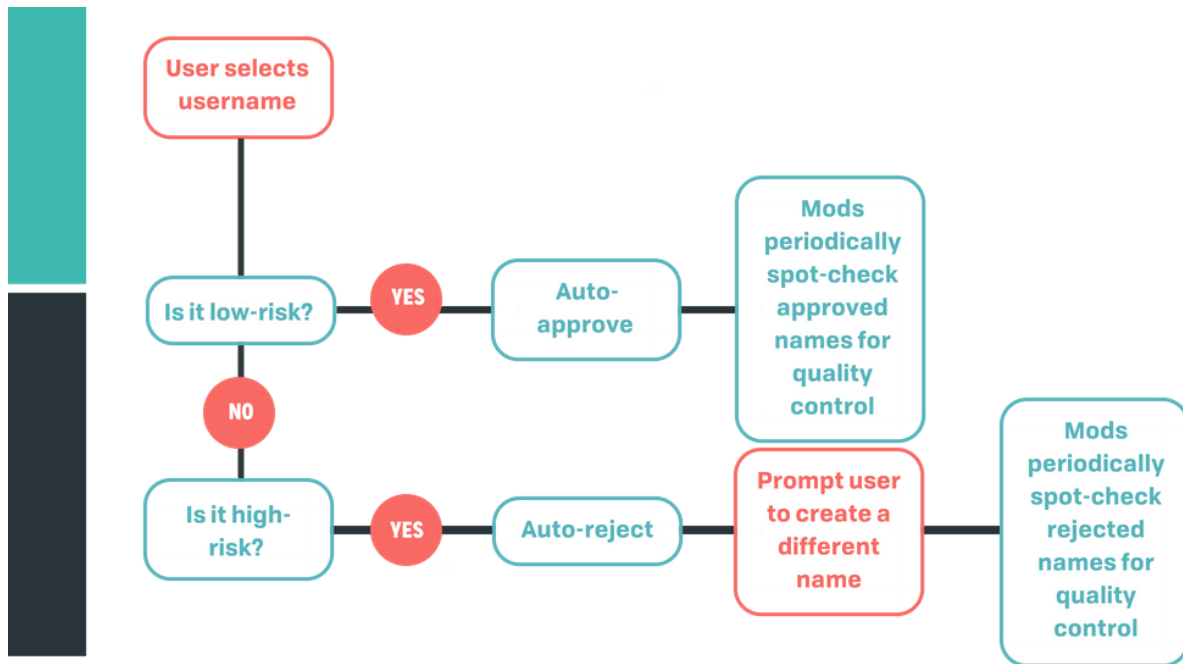
Accurately labelling PII can be challenging, so look for a content moderation solution that is equipped to identify personal information.

Username Workflow #1 - Over 13 Communities

For teen and adult communities, you can take a relatively simple approach to username moderation by automatically approving low-risk names and rejecting high-risk names.

Users can report medium-risk names, which moderators will review and approve or reject based on your Code of Conduct.

This workflow ensures that the moderation workload is manageable, with moderators focused on making nuanced decisions that AI can't handle.

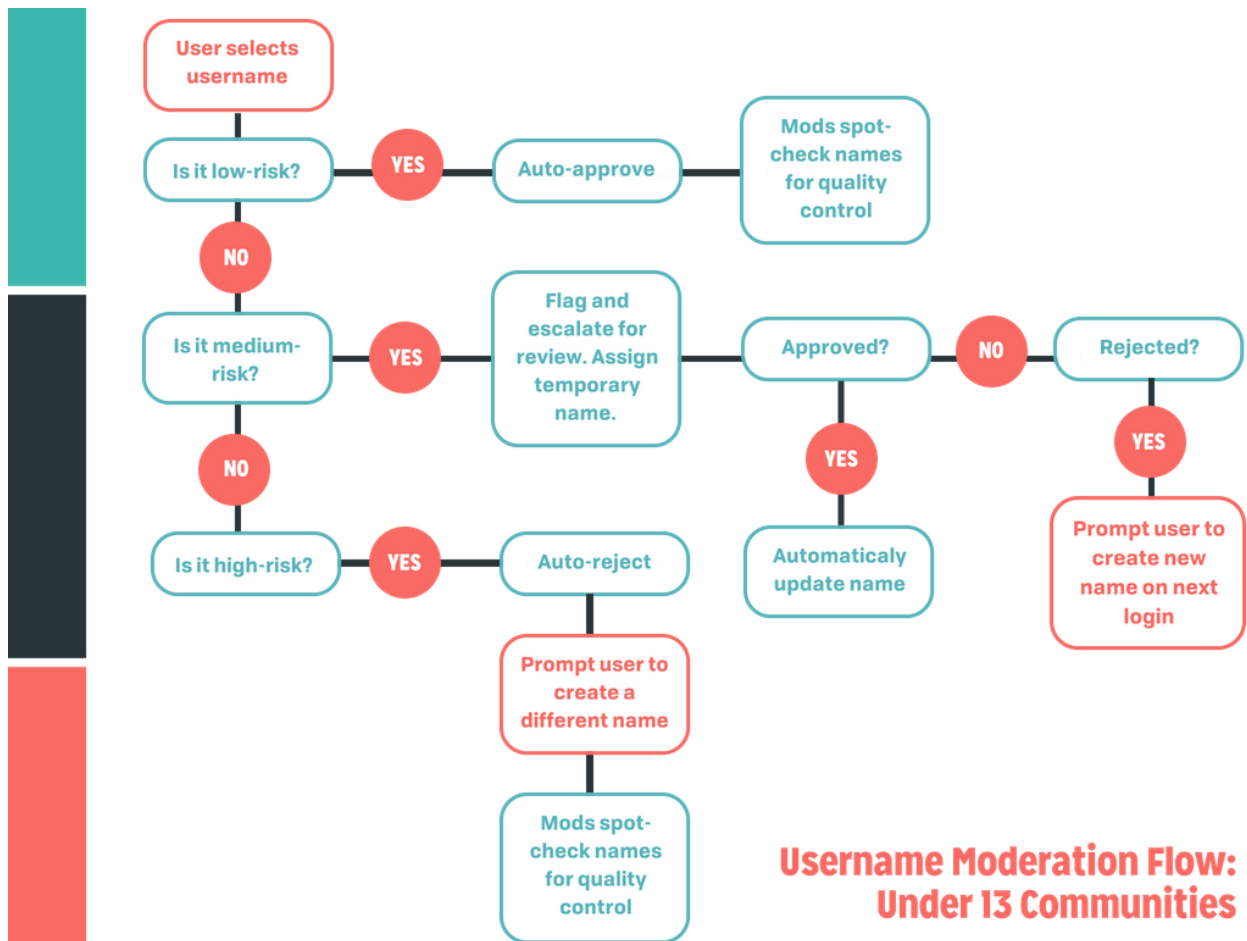


Username Moderation Flow: Over 13 Communities

Username Workflow #2 - Under 13 Communities

A slightly more refined approach is necessary for younger communities, especially if you're COPPA compliant.

This workflow finds a balance between safety, user experience, and moderation workload:



**Username Moderation Flow:
Under 13 Communities**

Quick Takeaways



A solid username moderation strategy is critical to community health and user retention.



Choose a content moderation tool that can handle the unique challenges of labeling, filtering, and flagging usernames.



Usernames fall under COPPA.



Tailor your moderation workflow to your community demographic.