

Moderation

This chapter covers general moderation tips, specific situations, and established practices within TNW.

- Getting Started

Getting Started

Moderation is a complicated subject with a variety of topics. This guide will have a few cheat sheets for specific scenarios, but most cases are complex and require you to develop different skills. By the end of this chapter, you'll have the basic tools to begin developing these skills. Don't be upset if you don't get it right the first time or if you need help from another staff member. As we'll discuss later, collaboration is key and you should allow others to help and guide you.

Code of Conduct

It is important to remember why moderation is required in online spaces. Most online social spaces have some kind of code of conduct to follow, and members are required to follow these rules. These rules can be split into two categories.

Systematic Rules

Systematic rules are rules put in place to inform users how to act in accordance with standard procedures. For example, a support server may require tickets to be filed in a specific way for it to be considered valid. Preferably, these types of rules should be implemented inherently into systems using permissions and design. However, this is not always possible or realistic, and so the rules must be communicated and consciously noted by users. One systematic rule we have on TNW is to keep all content in English.

When a user breaks a systematic rule, it usually is not a severe or hostile incident. Most often, systematic rules are broken out of forgetfulness or on accident, and their consequences are usually mild.

Behavioral Rules

Behavioral rules are the rules we'll be focusing most heavily on in this chapter because they are the most important to a community's health. Behavioral rules define how users should behave towards one another in the space.

A community's health can be measured by the friendliness and kindness of the users within the community. A healthy community fosters an environment where people can feel safe and trust each other.

In an ideal world, an online space would only need one behavioral rule: to be kind and courteous to others. However, differing life experiences result in different ideas of what "kind and courteous"

means, which can result in conflict. Establishing a more precise set of behavioral rules that defines the expectations allows all users to agree upon common ground.

It is crucially important when developing behavioral rules to remember that their purpose is to keep the community healthy. Rules should be targeted at minimizing hostility and ignorance and promoting kindness and understanding.

A common pitfall that many communities fall into is building behavioral rules that work like systematic rules. If behavioral rules are in set up with very specific boundaries that are more technically defined than behaviorally defined, the handling of issues will be algorithmic. This can cause issues where well-meaning users can be punished even if they haven't hurt anyone, and worse, malicious users will seek out exploits in the rule set as if it were a game. This situation moves the community *further* from being healthy, so it's important to make sure that behavioral rules are formed in such a way that they could not simply be automated by a computer.

For this reason, TNW's first rule is the most important rule that a community should have. It will be referred to as the **principle rule**.

“ 1. **Have a good attitude.**

- Treat everyone with respect and politeness. Being disruptive, aggressive, disrespectful, creepy, or incitive is prohibited.
- Hate speech of any kind is strictly prohibited.
- Threats toward the server are strictly prohibited.
- Fighting, trolling, spamming, instigating, and harassment is strictly prohibited.

This rule encompasses the purpose of behavioral rules, and you can typically decipher which rules are behavioral by asking if breaking them also violates the principle rule.

Key Ideas

1. The purpose of moderation is to uphold the code of conduct and maintain a healthy community.
2. Moderation is a complex task that requires skilled human judgement and does not serve its purpose if driven algorithmically.
3. Rules can be categorized as systematic rules or behavioral rules.
4. The purpose of systematic rules is to establish regular procedures.
5. The purpose of behavioral rules is to promote an agreeable expectation of kindness and understanding.