

Human-Machine Collaboration for Content GEORGIA TECH Regulation: The Case of Reddit Automoderator

Shagun Jhaver, Iris Birman, Eric Gilbert, Amy Bruckman

Georgia Institute of Technology

Summary

Research Questions

- 1. How are automated tools for content regulation created, acquired and used by Reddit moderators?
- 2. What are the benefits and challenges of using currently available automated regulation tools?
- 3. How does the use of automated mechanisms change the work of content regulation?



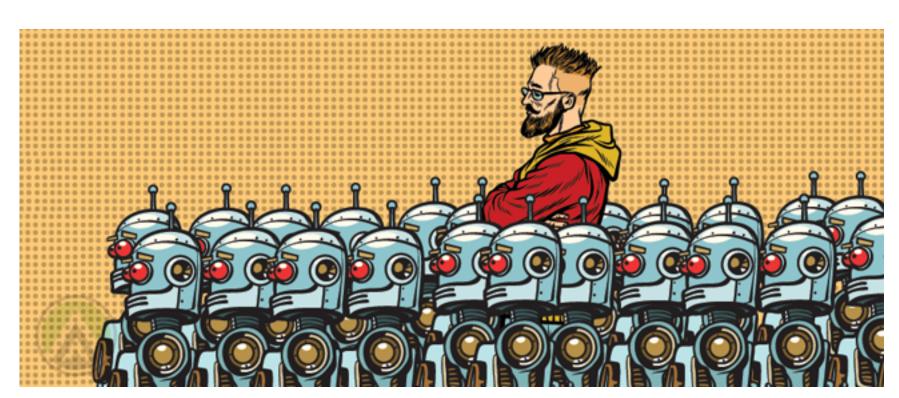
[-] AutoModerator image [M] 1 point 1 year ago

Your submission has been automatically removed because titles must begin with "PsBattle: ". This needs to be exact. Please refer to the Submission rules as to why this is the case and check to see that any other rules are satisfied if you choose to resubmit.

I am a bot, and this action was performed automatically. Please contact the moderators of this subreddit if you have any questions or concerns. permalink embed save give gold

Contributions

- 1. Understanding the content regulation ecosystem on Reddit.
- 2. Highlighting the additional tasks and challenges that using automated moderation tools introduce
- 3. Presenting design suggestions for mixedinitiative moderation systems.



Approach

- 1. Semi-structured Interviews with three moderators from each of the following subreddits:
- r/politics
- r/photoshopbattles
- r/space
- r/explainlikeimfive
- r/oddlysatisfying
- And Automod creator
- 2. Participant Observations
- 3. Inductive analysis of interview data.

Findings

• Introduction of Automod on Reddit

- 1. Moderators use a wide variety of automated tools, largely developed by volunteer users.
- 2. Automod was developed by Chad Birch as a moderation bot in 2012.
- 3. Reddit incorporated Automod and provide it as a private wiki to moderators on each subreddit.
- 4. Each Automod rule is written in YAML format:

#Remove comments for users with comment karma lower than 2 type: comment author: flair text (regex): "^\$" comment karma: "< 1" is submitter: false action: spam action reason: Low Karma

5. Automod has become an integral and indispensable part of Reddit.

Using Automod to Enforce Guidelines

- 1. Automod is well adept at enforcing some guidelines but not others.
- 2. Mods configure Automod to make easy calls but make decisions that require subjective interpretation themselves.



Need for Careful Curation of Automod

- 1. Broad rules can result in false positives.
- 2. Lack of information about Automod behavior
- 3. Reliance on user reports to find mistakes



Automod Creates New Tasks

- 1. Regular updating of Automod rules.
- 2. Preventing users from circumventing Automod.
- 3. Correcting false positives.