

Auditing Automated Content Moderation Tools

ANONYMOUS AUTHOR(S)*

We describe and audit content moderation tools provided by common platforms such as Discord and Twitch.

ACM Reference Format:

Anonymous Author(s). 2024. Auditing Automated Content Moderation Tools . 1, 1 (July 2024), 1 page. <https://doi.org/10.1145/nnnnnnnn>.
nnnnnnnn

1 INTRODUCTION

2 BACKGROUND

3 METHODOLOGY

4 EXPERIMENT DESIGN

5 RESULTS

6 DISCUSSION

REFERENCES

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM