Study | 03.05.2021

Reimagining content moderation and safeguarding fundamental rights

A study on community-led platforms

A study by Ben Wagner, Johanne Kübler, Eliška Pírková, Rita Gsenger, Carolina Ferro

INTRODUCTION

As millions of people log into social media platforms like Facebook, Twitter, and Reddit every day, the rules by which these platforms are governed increasingly determine how we interact with each other, and they shape the possibilities and nature of public discourse (Suzor 2020). The kinds of rules these platforms adopt, the policies they choose to enact, and the design choices they make all affect which information is available and how people communicate.

In line with the steep increase in user numbers and data generated per day, technology companies had to develop procedures to process a previously inconceivable amount of data. For instance, Facebook generates an estimated 4 petabytes of data every single day (Osman 2021). Responding to the need to examine and curate a large amount of data, platforms have developed content governance models and complex, multi-layered content moderation systems, which rely heavily on the removal of harmful and, otherwise, undesirable content.

However, there are growing concerns regarding the impact of those platforms' decisions on the freedom of expression and information and the digital rights of individuals. The focus on the blocking and deletion of content is accentuated through legislative approaches that also focus on deletion. In recent years, increased governmental pressure on online platforms "to do more" about the spread of hate speech, disinformation, and other societal phenomena online has led to a frenetic regulatory process across the European Union (EU), which, consequently, triggered similar regulatory responses around the globe. Due to the lack of

legal certainty in combination with unduly short time frames for content removal and the threat of heavy fines for non-compliance, platforms frequently over comply with these demands and swiftly remove large amounts of online content with no transparency and public scrutiny (Dara 2011; Ahlert, Marsden, and Yung 2004; Leyden 2004). The sheer volume of requests inevitably leads to erroneous takedowns, resulting in chilling effects for users faced with them (Penney 2019; Matias et al. 2020).

Many community-led platforms1 offer alternatives to these challenges for human rights and freedom of expression. However, these innovative approaches are typically not implemented by larger platforms. The alternative approaches often focus on community building and curation to strengthen communities to the point that content moderation is considerably less necessary. To accurately assess these alternative approaches, it is important to closely analyse the effects of different types of content moderation on user behaviour and their digital rights. Online communities without any content moderation at all are equally problematic for digital rights and typically quickly descend into what Daphne Keller termed the freedom of expression 'mosh pit.' (The Verge, 2021) Such online communities are not ideal for any of the actors involved, as only the loudest actors' voices can be heard.

This study explores alternative approaches to content moderation and, overall, different content governance models. Based on the research outcomes, it provides a set of recommendations for community-based and user-centric content moderation models that meet the criteria of meaningful transparency and are in line with international human rights frameworks. These recommendations are specifically addressed to EU lawmakers with the goal of informing the ongoing debate on the proposed EU Digital Services Act.

READ THE FULL STUDY

Recommended

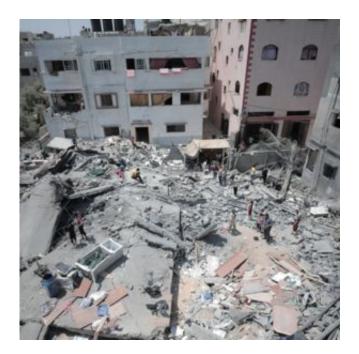Press release

Bild von Fabian Holtappels auf Pixabay

[Quote from MEP Alexandra Geese](#)

03.06.2025

Press release

Photo by Mohammed Ibrahim on Unsplash

[Greens/EFA call for suspension of EU-Israel Associatio…](#)

21.05.2025

News

© jonathan-marchal/unsplash

[Debriefing of the April 2025 Plenary Session](#)

04.04.2025

Press release

https://unsplash.com/photos/a-flag-on-a-building-_VE9mUHVd2s

[Commission's blank check to autocratic Egypt undermine…](#)

01.04.2025

Contact person

Narmine Abou Bakari

Digital Rights campaigner

Attached documents

[Study on community led online platforms](#)

Please share