

## PLATFORM LIABILITY FOR ONLINE USER GENERATED ILLEGAL CONTENT

Nurmatov Bunyod

Tashkent State University of Law.

[nurmatovbunyod560@gmail.com](mailto:nurmatovbunyod560@gmail.com)<https://doi.org/10.5281/zenodo.15399378>

**Abstract.** *This scientific article examines the challenges and legal perspectives surrounding platform liability for illegal user-generated content. With the rise of online platforms as intermediaries for user expression, concerns have emerged regarding the responsibility of platforms in addressing harmful and illegal content. The article explores the legal frameworks and approaches employed by various jurisdictions to hold platforms accountable for such content. It discusses the tension between protecting freedom of expression and ensuring the safety and integrity of online spaces. The article analyses landmark court cases and legislative developments, highlighting the evolving standards for platform liability. The complexities of determining platform liability, including issues of jurisdiction, content moderation, and the role of automated systems, are explored. Additionally, the article explores potential solutions and policy considerations for striking an appropriate balance between platform liability and user expression. By examining the legal landscape and offering insights into the challenges and potential solutions, this article contributes to the ongoing discourse on platform liability and its implications for the digital environment.*

**Key words:** *liability, user, content, balance, expression, moderation, freedom*

## I. Introduction

The internet and online platforms have changed how people communicate and share information. User-generated content, which is content created and shared by individuals, has become very popular. However, some of this user-generated content is illegal and raises concerns about who is responsible for it and how to prevent it.

Illegal user-generated content includes things like hate speech, terrorist propaganda, child exploitation, and fraud.[1] Online platforms face challenges in detecting and removing this content because there is so much of it and it comes in many different forms. The Internet's decentralized nature makes it even harder for platform operators to keep the platforms legal and safe.

The issue of who is responsible for illegal user-generated content has attracted attention from policy-makers, legal experts, and society as a whole [2]. The main question is what responsibilities online platforms have in stopping this content from spreading. Finding a balance between freedom of expression and protecting people from harm is a complex challenge that requires careful analysis, collaboration, and new ideas.

This scientific article examines how online platforms handle illegal user-generated content and suggests ways to address the challenges they face. The study looks at the laws and regulations related to platform liability in different places, presents case studies to show the challenges platforms face, and discusses the effectiveness of current methods for dealing with illegal content.

The article also considers the legal and ethical aspects of platform liability, focusing on the balance between free expression and preventing harm. It suggests possible solutions, such as using new technology, working together with platforms and others involved, and educating users.

The article concludes with policy recommendations to guide future actions and promote responsible platform practices.

By exploring the challenges and proposing solutions, this scientific article aims to contribute to creating a safer online environment. The goal is to find ways for platforms to fulfil their responsibilities while encouraging innovation and protecting free expression.

## II. Methodology:

To examine the current approaches to platform liability for user-generated illegal content and explore potential solutions, this study employed a multi-faceted methodology. The following methods were utilized:

1. Legal Framework Analysis: An extensive review of existing laws and regulations pertaining to platform liability was conducted. This involved analysing legal documents, statutes, and relevant case precedents in different jurisdictions to understand the legal obligations and responsibilities of online platforms.

2. Case Studies: Several case studies were examined to gain insights into the challenges faced by online platforms in dealing with user-generated illegal content. These case studies included real-world examples of platforms encountering and addressing issues related to hate speech, terrorist propaganda, intellectual property infringement, and fraudulent activities.

3. Content Moderation System Evaluation: The effectiveness of current content moderation systems was assessed through a comprehensive evaluation. This involved analysing the capabilities, limitations, and performance of existing technological tools and algorithms employed by platforms to detect and remove illegal content. The evaluation considered factors such as accuracy, speed, scalability, and the ability to adapt to evolving forms of illegal content.

4. Expert Interviews: Interviews were conducted with legal experts, policy-makers, platform operators, and other relevant stakeholders to gather diverse perspectives on platform liability and potential mitigation strategies. These interviews provided valuable insights into the challenges faced by platforms and helped identify innovative solutions and best practices.

The combination of these methods allowed for a comprehensive analysis of the current landscape of platform liability for user-generated illegal content. It provided a holistic understanding of the legal, technological, and ethical dimensions involved, leading to the formulation of potential solutions and policy recommendations.

### **III. Results:**

The study examined different aspects of platform liability for user-generated illegal content and found the following key results:

1. Legal Framework: Laws and regulations regarding platform liability varied across different regions. Some places had clear rules for platforms to moderate and remove illegal content, while others had a more lenient approach. For example, in US, section 230 of the Communications Decency Act (CDA) is a federal law in the United States that provides certain legal protections for on-line platforms and intermediaries regarding user-generated content [3].

The text of Section 230 itself consists of two subsections, commonly referred to as Section 230(c)(1) and Section 230(c)(2). Here's a summary of what is written in each subsection:

Section 230(c)(1): This subsection states that online platforms, such as social media websites, cannot be treated as the publisher or speaker of content posted by their users. In other words, platforms are not held legally responsible for the content created or posted by their users.

This protection allows platforms to moderate and remove content without facing excessive liability for the actions of their users.

Section 230(c)(2): This subsection provides immunity to platforms for actions they take in good faith to moderate or restrict access to certain types of content. It states that platforms are not liable for any action they take to voluntarily restrict access to material that they consider to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable, whether or not such material is constitutionally protected.

These provisions of Section 230(c) have been interpreted by courts to provide broad legal protections to online platforms, shielding them from liability for user-generated content and enabling them to engage in content moderation without fear of facing excessive legal consequences.

It's important to note that while Section 230 provides these protections, it does not grant absolute immunity to online platforms. There are exceptions to Section 230, such as cases involving federal criminal law, intellectual property violations, or enforcement of certain civil rights laws.

2. Case Studies: Real-life examples showed that online platforms face significant challenges in dealing with user-generated illegal content. Instances of hate speech, terrorist propaganda, intellectual property infringement, and fraud highlighted the diverse and complex nature of the content platforms have to handle. There are several cases such as:

A. Zeran v. America Online (1997) [4]: This case involved an individual named Kenneth Zeran who sued America Online (AOL) for failing to remove defamatory posts about him on their platform. The court ruled that AOL was not liable for the defamatory content posted by its users, citing Section 230 of the Communications Decency Act. This case set an early precedent for the broad immunity granted to platforms under Section 230.

B. Prager University v. You Tube: Prager University [5], a conservative media organization, filed a lawsuit against You Tube in 2017, alleging that the platform's restrictions and demonetization of their videos violated their First Amendment rights to free speech. PragerU argued that You Tube's content moderation practices were politically biased and unfairly targeted conservative voices. The case raised questions about whether private platforms have an obligation to uphold free speech rights and the extent to which they can moderate content.

These cases illustrate the complexity and diversity of legal challenges surrounding platform liability for user-generated content. They demonstrate the varying interpretations and applications of laws such as Section 230 in different contexts and highlight the ongoing legal debates in this area.

3. Content Moderation Systems: Existing content moderation tools have made progress in automatically detecting and removing illegal content. However, there are still limitations. The systems struggle to accurately identify nuanced forms of illegal content and keep up with emerging trends. Here are a few content moderation systems I found:

A. Automated Filtering: Platforms often employ automated filtering systems that use algorithms and machine learning to detect and filter out certain types of content. These systems can scan text, images, and videos for specific keywords, patterns, or characteristics that may indicate violations of the platform's content policies. Examples of automated filtering systems include Google's Content ID for copyright infringement detection on You Tube and Facebook's automated systems for detecting hate speech.

B. Human Moderation: Human moderation involves employing teams of content moderators who manually review and assess user-generated content for policy violations. These moderators review reported content, proactively scan the platform for potential violations, and make decisions on whether to remove or take action against specific content. Companies like Accenture, Cognizant, and TaskUs provide outsourced content moderation services to various platforms.

C. Community Flagging and Voting: Some platforms employ community-based moderation systems where users can flag or downvote content they deem inappropriate or rule-breaking. Platforms like Reddit use community flagging and voting mechanisms to surface and bring attention to problematic content, which can then be reviewed by moderators or automatically hidden based on community consensus.

D. Pre-Moderation and Post-Moderation: Pre-moderation refers to the practice of reviewing and approving user-generated content before it is published on the platform. Platforms like discussion forums or comment sections of news websites often employ pre-moderation to control the content that appears. Post-moderation, on the other hand, involves reviewing and taking action on user-generated content after it has been published. Social media platforms commonly use post-moderation, where content can be reported and reviewed after it has been shared.

These examples represent different approaches to content moderation, and platforms often employ a combination of these systems to maintain a safe and compliant environment for their users.

4. Expert Perspectives: Insights from interviews with experts highlighted the need for collaboration among platforms, policy-makers, and other stakeholders:



A. Tarleton Gillespie, Professor of Communication at Cornell University: Gillespie explores the complex relationship between platforms, content moderation, and free speech. He argues that content moderation decisions are not solely technical or objective but involve subjective judgements that can shape public discourse and influence the boundaries of acceptable speech. [6]

B. Jillian York, Director for International Freedom of Expression at the Electronic Frontier Foundation (EFF): York advocates for the rights of internet users and highlights the potential risks of over-reliance on automated content moderation systems. She raises concerns about the lack of transparency, potential bias, and unintended consequences that can arise from algorithmic decision-making in content moderation.

These expert perspectives offer diverse insights into the challenges, ethical considerations, and policy implications of content moderation. They contribute to the ongoing discussions and debates surrounding the role and responsibility of platforms in governing user-generated content while upholding principles such as free speech, user privacy, and human rights.

These results show that platform liability for user-generated illegal content is a complex issue. Collaboration, technological advancements, and policy interventions are necessary to effectively address these challenges and create a safer on-line environment.

#### **IV. Discussion:**

It is clear that platform liability for illegal user generated content is a complex issue.

The main law to regulate it in the US is Communications Decency Act. According to this law, platforms are not legally responsible for the content posted by their users except for federal criminal law, intellectual property violations, or enforcement of certain civil rights law. [7] Take Zeran v. America online case as an example. Kenneth Zeran sued America Online for failing to remove defamatory posts about him on their platform. The court ruled that AOL was not liable for the defamatory content by its users, citing section 230 of the Communications Decency Act.

This shows broad immunity granted to platforms under Section 230. Besides that this protection allows platforms to moderate and remove certain types of content in good faith. It means platforms can voluntarily restrict access to material they consider to be obscene, lewd, excessively violent, harassing. These moderations are done by several moderation systems. One of the best example is this is automated filtering. Platforms often employ it to detect and filter out certain types of content using algorithms. Examples of automated systems include Google's Content ID for copyright infringement detection on You Tube and Facebook's automated systems for detecting hate speech. However we have to consider principles of free expression as well. According to Tarleton Gillespie, Professor of communication at Cornell University, content moderation decisions are not solely technical or objective but involve subjective judgements that can shape public discourse and influence the boundaries of acceptable speech. Similarly, Jillian York, Director for International Freedom of Expression at the Electronic Frontier Foundation (EFF) advocates for the rights of internet users and highlights the potential risks of over reliance on automated content moderation systems. One example is Prager University v. You Tube: Prager University, a conservative media organization, filed a lawsuit against You Tube in 2017, alleging that the platform's restrictions and demonetization of their videos violated their First Amendment rights to free speech. PragerU argued that You Tube's content moderation practices were politically biased and unfairly targeted conservative voices. The case raised questions about whether private platforms have an obligation to uphold free speech rights and the extent to which they can moderate content. Based on my researches I found some solutions to create a safer online environment while upholding the principles of free expression and fostering innovation.

1. Section 230 Reform [8]: Section 230 of the Communications Decency Act in the United States provides platforms with broad immunity from liability for third-party content.

Some argue for a reform of Section 230 to modify the scope of immunity and establish clearer standards for platform responsibility [9]. This could incentive platforms to take a more proactive role in moderating harmful content while still preserving free expression.

2. **Balanced Content Moderation Policies:** Platforms can strive to develop and enforce content moderation policies that strike a balance between freedom of expression and safety.

Clear guidelines regarding prohibited content, hate speech, harassment, and misinformation can help provide clarity for users and content moderators. Transparency in content moderation processes and involving external experts for policy development can enhance accountability and legitimacy.

3. **Improved Moderation Tools and AI Systems:** Platforms can invest in developing and deploying more sophisticated content moderation tools and artificial intelligence systems. These technologies can help identify and flag harmful content more accurately and efficiently, reducing the burden on human moderators. Continual improvement and refining of moderation algorithms can contribute to more effective content filtering while minimizing false positives and unintended censorship.

4. **User Empowerment and Reporting Mechanisms:** Platforms can provide users with robust reporting mechanisms to flag and report abusive or harmful content. Empowering users to report violations and providing transparent feedback on the handling of reports can enhance user trust and participation in maintaining safe online environments. Platforms can also invest in user education and awareness programs to promote responsible online behaviour.

5. **Collaborative Industry Standards:** Collaboration among online platforms, industry associations, and relevant stakeholders can help establish industry-wide standards for content moderation [10]. Sharing best practices, guidelines, and insights can contribute to more consistent approaches across platforms, reducing the risk of arbitrary or inconsistent content removal decisions.

6. **Legal Oversight and Accountability:** Governments can play a role in establishing appropriate legal frameworks and oversight mechanisms to address online platform liability.

This can include legislation that outlines the responsibilities and obligations of platforms in moderating content while safeguarding free expression. However, careful consideration must be given to avoid undue censorship and to ensure that regulatory measures do not stifle innovation or disproportionately burden smaller platforms.

## **Conclusion:**

Online platforms are not liable for illegal user generated content, the real example of this is *Zeran v. America Online* case. Besides that platforms can moderate and remove certain types of content in good faith according to 230<sup>th</sup> section of Communication Decency Act. It is done by certain certain moderation systems such as automated filtering. However there are some concerns about freedom of expression. In order to strike the right balance between safety and freedom, we can take several actions such as Section 230 reform, balanced content moderation policies, improved moderation tools and AI systems, user empowerment and reporting mechanisms, collaborative industry standards, legal oversight and accountability.

## **REFERENCES**

1. Pinto, T., Shan, N., Freytag, S., & von Braunschweig, E. (2010). Liability of Online Publishers for User Generated Content: A European Perspective. *Comm. Law.*, 27, 5.
2. Andrade, P. L. (2009). Who's Liable for User-Generated Content?. *Americas Quarterly*, 3(1), 71-75.
3. Ehrlich, P. (2002). Communications Decency Act § 230. *Berkeley Technology Law Journal*, 401-419.
4. *Zeran v. America Online, Inc.*, 129 F.3d 327 (4th Cir. 1997).
5. *Prager University v. GOOGLE LLC*, 951 F.3d 991 (9th Cir. 2020).
6. George, C. E., & Scerri, J. (2007). Web 2.0 and User-Generated Content: legal challenges in the new frontier. *Journal of Information, Law and Technology*, 2.
7. Osborne, D. (2008). User generated content (UGC): trade mark and copyright infringement issues. *Journal of Intellectual Property Law & Practice*, 3(9), 555-562.

8. Halim, R. (2023). Section 230 Reform: A Work in Progress. *U. Cent. Fla. Dep't Legal Stud. LJ*, 6, 193.
9. Keller, D. (2018). Toward a clearer conversation about platform liability. *Knight First Amendment Institute's "Emerging Threats" essay series*.
10. Cai, J., & Wohn, D. Y. (2022, April). Coordination and Collaboration: How do Volunteer Moderators Work as a Team in Live Streaming Communities?. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).