NLUA Journal of Intellectual Property Rights ISSN: 2583-8121 (Online)
Volume 3 Issue 2

ALGORITHMIC CONTENT MODERATION, IP AND EXPRESSION: NAVIGATING THE TENSION THROUGH THEORIES

Rutambhara Nayak[*]

Anindya Sircar[**]

Abstract

The fast-paced evolution of digital technologies and the rise of algorithmic content moderation have altered the dynamics between participatory culture and Intellectual Property (IP) enforcement. This paper analyses the broader effects of Algorithmic content moderation on participatory culture, particularly in the realm of IP enforcement. It explores how these algorithmic content moderation systems are designed to manage IP rights and the resulting impact on creativity, innovation, and the circulation of cultural materials. Furthermore, this study delves into the perspectives of various stakeholders (rights holders, creators, and users) who must maneuver the complexities and challenges brought about by these automated systems. By addressing these aspects, the paper aims to bring forth the inherent tensions between IP enforcement and cultural expression, providing insights into the consequences of automated decision-making in today's digital world.

Keywords: Algorithmic Content Moderation, Intellectual Property, Participatory Culture, Theories

1. Introduction

Legal automation is possible thanks to advances in artificial intelligence, machine learning, computing power, and big data analytics, but they also threaten civil freedoms and cultural values. This research is limited and lacks clarification on several crucial issues, including the necessity for detailed empirical studies on how legal automation may damage privacy and individual liberties.[1] Instagram, TikTok, Netflix, Swiggy, and Uber have greatly impacted the realisation of crucial public ideals and policy

[*] Research Assistant, DPIIT-IPR Chair, NALSAR University of Law, Hyderabad.
[**] Chair Professor – IPR, DPIIT IPR Chair, NALSAR University of Law, Hyderabad.
[1] Jonathon W. Penney, "Privacy and Legal Automation: The DMCA as a Case Study" 22 *Stanford Technology Law Review* 412 (2019).

objectives related to media, civic participation, education, and transport.[2] Online platforms have increasingly used algorithmic law enforcement. These intermediaries distribute content and are suitable partners for civil and criminal enforcement since they can monitor and block access to illicit information. Copyright law has led this change. Under the Digital Millennium Copyright Act, platforms that remove allegedly infringing content when contacted have safe harbor protection since the early 1990s. Over the past two decades, prominent platforms have integrated the Notice and Takedown (N&TD) process. Copyright holders increasingly employ automated systems to send enormous numbers of takedown requests, while online intermediaries use algorithms to filter, block, or remove content with little human intervention.[3] This tug of war has created a worrisome situation that may be resolved with a balanced approach.

1.1.Theoretical Framework

This sub-section elaborates upon the theories of Intellectual Property and participatory culture respectively for understanding the jurisprudence behind the both and in return, trace their interaction in the digital world to their theoretical foundation.

1.1.1. Legal Theories on Intellectual Properties

Legal theories on Intellectual Property, primarily provide the framework for understanding the balance between protecting creators' rights and fostering cultural exchange and subsequently these theories also catalyse useful conversations among the various people and institutions responsible for the shaping of the law.[4] The primary theories include:

Natural Rights Theory (Lockean Theory)

Propounded on John Locke's views, Natural Rights theory holds that creators have the right to control and use their Intellectual Property because it represents their effort and uniqueness.[5] Locke argues that blending labor with basic materials creates

[2] Natali Helberger, Jo Pierson and Thomas Poell, "Governing Online Platforms: From Contested to Cooperative Responsibility," 34 *Information Society* 1–14 (2018).

[3] Maayan Perel and Niva Elkin-Koren, "Accountability in Algorithmic Enforcement: Lessons from Copyright Enforcement by Online Intermediaries", 19 *Stanford Technology Law Review* 473 (2016).

[4] William Fisher, "Theories of Intellectual Property", in Stephen R. Munzer (ed.), *New Essays in the Legal and Political Theory of Property* 168-199 (Cambridge University Press, 2001).

[5] Justin Hughes, "The Philosophy of Intellectual Property," 77 *Georgetown Law Journal* 287 (1988).

moral and legal worth, including intellectual creations.[6] Critics say applying this paradigm to IP may overprotect and limit knowledge and cultural resources.[7]

Utilitarian Theory

This theory holds that IP rights compensate innovators economically for their creativity and innovation. This utilitarian method weighs the benefits of new works against their societal costs, such as higher consumer prices, to optimise welfare. Intellectual property policy aims to protect valued creative works while minimising social cost.[8]

Labour Theory

This widely accepted theoretical approach holds that when someone works hard on unowned or shared resources, they inevitably gain the right to what they produce. Lockean ideas in Intellectual Property views "facts and ideas" as shared building blocks, and the labor involved in converting them is vital to the ultimate product.[9]

Personality Theory (Hegel's Theory)

This approach emphasises the personal relationship between artists and their creations, pushing for moral rights that protect their work's financial and personal integrity. It cites Kant and Hegel to argue that Intellectual Property Rights protect artists' expression, which defines them. Unauthorised changes or misuse of their work would violate this personal connection.[10]

Social Planning Theory

Intellectual Property Rights help achieve social goals like public welfare, cultural variety, and knowledge availability. Social planning theory promotes cultural exchange and creation while preserving IP. Jefferson, the early Marx, the Legal Realists, and ancient and modern champions inspire this type of theorist. This theory shares

---

6 Tom Palmer, "Are Patents and Copyrights Morally Justified? The Philosophy of Property Rights and Ideal Objects," 13 *Harvard Journal of Law & Public Policy* 817 (2005).

7 Jeremy Waldron, "From Authors to Copiers: Individual Rights and Social Values in Intellectual Property," 68 *Chicago-Kent Law Review* 841 (1992).

8 Robert P. Merges, "Introduction: Main Themes" in *Justifying Intellectual Property* 1–28 (Harvard University Press, 2011).

9 *Supra* note 5 at 289.

10 *Ibid.*

utilitarianism's focus on accomplishing societal goals, but it embraces a richer vision of an ideal society beyond "social welfare".[11]

Commons-Based Approach

This strategy promotes knowledge sharing and collaborative production, undermining intellectual property laws. Both commons and exclusive rights stimulate creation and discovery, therefore Intellectual Property laws should balance both when managing works and innovations. Julie Cohen[12] suggests revising copyright to combine private and public rights into a more complex and dynamic mix of rights and privileges.

Prospect Theory (Kitch's Theory)

Prospect Theory, as explained by Kitch,[13] states that early IP rights give incentives for subsequent development and commercialisation. IP rights are similar to mining claims, according to Kitch,[14] and early allocation of rights helps manage resources and stimulate innovation by protecting R&D investments. This idea has supported broad patent rights, especially in pharmaceuticals, where product development requires major expenditure.[15]

Schumpeterian Theory

Schumpeterian Theory, based on economist Joseph Schumpeter, advocates that entrepreneurs drive economic growth through "creative destruction", where new ideas and technologies replace old ones, according to Schumpeter. IP rights reward entrepreneurs for taking risks and provide financial incentives for innovation. The theory acknowledges that excessive protection could lead to monopolies, which could hamper innovation and healthy competition.[16]

Incentive Theory

Incentive Theory states that IP rights encourage creativity and innovation in individuals and firms. Without the promise of exclusive rights, the argument claims, there

---

11 *Supra* note 4.

12 Julie E. Cohen, "Copyright, Commodification, and Culture: Locating the Public Domain", in Lucie Guibault and P. Bernt Hugenholtz (eds.), *The Future of the Public Domain: Identifying the Commons in Information Law* 121–166 (Kluwer Law International, The Hague, 2006).

13 Edmund W. Kitch, "The Nature and Function of the Patent System", 20(2) *Journal of Law and Economics* 265–290 (1977).

14 *Ibid*

15 Mark A. Lemley, "Property, Intellectual Property, and Free Riding", 83 *Texas Law Review* 1040 (2005).

16 Joseph A. Schumpeter, *Capitalism, Socialism and Democracy* (Harper & Brothers, New York, 1942).

would be little motivation to devote time and money into developing new ideas, as others might freely duplicate and profit from them.[17] The difficulty is combining incentives with IP rights that do not impede access and innovation.

Economic Theory

Economic theory views IP through market efficiency and welfare maximisation. It claims that IP rights should fix market failures caused by intellectual goods' non-excludability and non-rivalry. IP rights can assist inventors recover their investments and foster socially important technologies by giving temporary monopolies.[18] This idea also advocates limiting IP rights' scope and term to avoid monopoly pricing and knowledge loss.[19]

Welfare Theory

IP Welfare Theory focuses on social well-being and suggests that IP protection should be combined with compulsory licensing or fair usage to prevent monopolies. Welfare Theory supports flexible IP rules that meet society requirements, especially in public health and education.[20]

Reward Theory

Reward Theory supports IP rights to reward creators for their social contributions. IP rights compensate artists for their hard work and creativity, according to this viewpoint.[21] This notion supports justice by arguing that individuals who spend time and money on new ideas should profit. Reward Theory must ensure that awards are proportionate and do not over control cultural and scientific resources.[22]

---

17 William M. Landes and Richard A. Posner, *The Economic Structure of Intellectual Property Law* 222 (Harvard University Press, Cambridge, 2003).

18 Kenneth J. Arrow, "Economic Welfare and the Allocation of Resources for Invention", in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, National Bureau Committee for Economic Research and Committee on Economic Growth of the Social Science Research Council 606-626 (Princeton University Press, Princeton, 1962).

19 *Supra* note 4.

20 Joseph E. Stiglitz, "Economic Foundations of Intellectual Property Rights", 57 *Duke Law Journal* 1693–1724 (2008).

21 Louis Kaplow, "The Patent-Antitrust Intersection: A Reappraisal", 97 *Harvard Law Review* 1813–1892 (1984).

22 Brett M. Frischmann, *Infrastructure: The Social Value of Shared Resources Infrastructure* (Oxford University Press, 2012).

'Race to Invent' Theory (Merges and Nelson)

The 'Race to Invent' Theory by Robert Merges and Richard Nelson examines innovation competition. It argues that IP rights spur inventors to explore and patent new ideas, accelerating technological progress.[23] This paradigm acknowledges pitfalls such patent thickets and strategic conduct that inhibit competition and innovation.[24] IP regulations should be carefully calibrated to foster healthy competition without introducing undue impediments to entry, according to the notion.

Rent Dissipation Theory (Grady and Alexander)

Mark Grady and Jay Alexander's Rent Dissipation Theory claims that IP rights can dissipate economic rents, or excess earnings from exclusive resource control. According to this idea, competitive IP rights pursuit can lead to over-investment in patent races, where patent-securing and enforcement costs outweigh benefits.[25] The argument proposes that IP laws should prioritise IP rights quality over quantity to reduce rent dissipation.

Public Domain Theory

A strong public domain where ideas and cultural resources are freely available is central to Public Domain Theory. This idea states that wide or long-term IP rights can degrade the public domain, reducing knowledge and creativity.[26] Public Domain Theory promotes IP rules that balance protection and accessibility to allow the public to use the cultural and scientific commons.

Cultural Theory

Cultural Theory addresses IP through cultural behaviors and creativity social processes and proposes that IP laws should reflect the varied ways individuals create and exchange cultural works rather than enforcing a commercial paradigm.[27] Cultural Theory

---

[23] Robert P. Merges and Richard R. Nelson, "On the Complex Economics of Patent Scope," 90 *Columbia Law Review* 839–916 (1990).

[24] Michael A. Heller and Rebecca S. Eisenberg, "Can Patents Deter Innovation? The Anticommons in Biomedical Research," 280(5364) *Science* 698–701 (1998).

[25] Mark F. Grady and Jay I. Alexander, "Patent Law and Rent Dissipation," 78 *Virginia Law Review* 305-350 (1992).

[26] James Boyle, *The Public Domain: Enclosing the Commons of the Mind* preface xv-xvi (Yale University Press, 2008).

[27] Peter Krapp, "The Cultural Life of Intellectual Properties: Authorship, Appropriation, and the Law (review)," 58 *Cultural Critique* 198–201 (2004).

emphasises conserving traditional knowledge, folklore, and communal innovation, which may not fit into IP frameworks.[28]

'Property begetting Property' Theory

The 'Property begetting Property' Theory states that IP rights tend to produce more rights, expanding property claims over time. This hypothesis shows that IP rights proliferate as rights holders seek to limit associated innovation or expression.[29] It argues that IP legislation should be properly handled to prevent rights overconcentration and allow new market entrants.[30]

Distributive Justice Theory

Distributive Justice Theory addresses IP from a justice and equality perspective in resource and benefit distribution. This idea promotes programs that increase knowledge and cultural access for marginalised and disadvantaged groups.[31] Distributive Justice Theory emphasises balancing creator and user interests to prevent IP rights from worsening social inequality.[32]

Morality Theory

Morality Theory argues that the creation and use of intellectual works should align with broader moral values, such as respect for human dignity, fairness, and justice.[33] It supports the idea that IP rights should not be absolute but should be tempered by considerations of the public good and the potential impact on human rights.[34] It advocates for IP laws that reflect ethical principles and promote the responsible use of creative and intellectual resources.[35]

---

28 Michael F. Brown, "Heritage Trouble: Recent Work on the Protection of Intangible Cultural Property", 12(1) *International Journal of Cultural Property* 40–61 (2005).

29 William W. Fisher III, "The Growth of Intellectual Property: A History of the Ownership of Ideas in the United States", in Hannes Siegrist and David Sugarman (eds.), *Eigentum im internationalen Vergleich* 265–290 (1999).

30 *Ibid*.

31 Leah Chan Grinvald, "Against Progress: The Value of Distributive Justice in Intellectual Property", 102 *Boston University Law Review Online* 74-75 (2022).

32 *Ibid*.

33 Adam D. Moore, "Intellectual Property, Innovation, and Social Progress: The Case Against Incentive Based Arguments," 26 *Hamline Law Review* 602–630 (2003).

34 Peter Drahos, *A Philosophy of Intellectual Property* (Dartmouth Publishing, Aldershot, 1996).

35 Richard A. Spinello, "The Future of Intellectual Property," 5 *Ethics and Information Technology* 1–16 (2003).

Critical Legal Studies (CLS) Theory

Critical Legal Studies Theory questions the power structures and interests that shape IP law. CLS scholars claim that IP law favours wealthy economic entities over the public.[36] This proposition shows how IP law can exacerbate social and economic inequality, especially in emerging nations.[37] CLS Theory promotes a more critical and inclusive IP law that analyses IP rights' social and political effects.[38]

Humanitarian Perspective

The Humanitarian Perspective on IP highlights IP laws' role in tackling global issues like healthcare, food security, and education. This view holds that IP rights should benefit everyone, especially developing nations that may be disproportionately harmed by restrictive IP regimes.[39] Mandatory licensing, patent pools, and open-access information exchange are supported by the Humanitarian Perspective.[40]

Each of these theories offers a different perspective on why Intellectual Property Rights exist, what their purpose should be, and how they should be structured. In practice, modern Intellectual Property law often reflects a combination of these theories, balancing the rights of creators with the needs of society.

1.1.2. Theories of Participatory Culture

Theories of participatory culture explore the dynamics of cultural production and consumption in a digital age where individuals are active participants in shaping cultural narratives. The concept of participatory culture has gained significant attention, particularly with the rise of digital media, where users are not just passive consumers but also active producers and sharers of content. Here are some key theories and concepts related to participatory culture:

---

[36] Mark Kelman, *A Guide to Critical Legal Studies* (Harvard University Press, Cambridge, MA, 1987).
[37] Amy Kapczynski, "The Access to Knowledge Mobilization and the New Politics of Intellectual Property," 117 *Yale Law Journal* 804–885 (2008).
[38] Mark Tushnet, "Critical Legal Studies: A Political History," 100(5) *Yale Law Journal* 1515–1544 (1991).
[39] Susan K Sell, "What Role for Humanitarian Intellectual Property? The Globalization of Intellectual Property Rights", 6(1) Minnesota Journal of Law, Science & Technology 191-192 (2004).
[40] *Supra* note 20.

Henry Jenkins' Theory of Participatory Culture

Henry Jenkins proposed convergence culture, which blends media platforms, content, and audiences. Participatory culture blurs the borders between producers and consumers, allowing audiences to collaborate and modify media content.[41] Jenkins emphasises that participatory culture is about community and collaboration as much as content creation and opposes top-down media creation, where a few create and many consume material. Instead, it encourages democratic, decentralised media production.

Collective Intelligence

Jenkins in this theory posits that participatory culture harnesses the collective knowledge and creativity of individuals, leading to new forms of problem-solving, innovation, and cultural production. The collaborative nature of participatory culture enables diverse perspectives to converge, enhancing the richness of cultural expression.[42]

Prosumer Theory

Prosumption involves both production and consumption rather than focusing on either one (production) or the other (consumption). This theory examines the role of individuals as both producers and consumers (prosumers) of content. It emphasises the agency of individuals in shaping cultural narratives and highlights the impact of digital technologies in empowering prosumers to participate actively in cultural production.[43]

Participatory Design

This theory explores the potential of participatory culture to enhance democratic engagement and civic participation. By providing individuals with platforms to express their voices, participatory culture can facilitate dialogue, social movements, and political activism.[44]

---

41 Henry Jenkins, *Convergence Culture: Where Old and New Media Collide* (New York University Press, 2006).

42 Henry Jenkins, *Fans, Bloggers, and Gamers: Exploring Participatory Culture* (New York University Press, 2006).

43 George Ritzer and Nathan Jurgenson, "Production, Consumption, Prosumption: The Nature of Capitalism in the Age of the Digital 'Prosumer'," 10 *Journal of Consumer Culture* 13–36 (2010).

44 Henry Jenkins, Mizuko Ito, *et.al*., *Participatory Culture in a Networked Era: A Conversation on Youth, Learning, Commerce, and Politics* (Polity Press, Cambridge, 2016).

Remix Culture

Participatory culture relies on remix culture, where people creatively remix, combine, or repurpose content. This culture, according to Lawrence Lessig[45] values the idea that media and cultural artifacts should be interacted with, reinterpreted, and reinvented. This approach typically generates legal issues, particularly copyright law, which can restrict free speech.

Fan Culture and Fandoms

Fandoms are another important feature of participatory culture where fans often create fan fiction, fan art, and other derivative works from media materials to build a community. These activities show a personal connection to the original content and a shared interest in its interpretation and expansion.[46] Fans claim their right to create and share cultural content, challenging authorship and ownership. Scholars say fandoms exist in a legal grey area where participatory culture and IP law clash.[47]

Transmedia Storytelling

Henry Jenkins proposed transmedia storytelling, which involves conveying a tale or message across many media platforms, with each platform contributing to the overall narrative.[48] Transmedia storytelling requires audience engagement to understand and expand the story across media. The collaborative character of transmedia storytelling requires a more complex view of IP that enables content sharing.[49]

Civic Engagement and Social Movements

Participatory culture is tied to civic involvement and social movements, which use media for activism and public discourse. Social media platforms have accelerated information dissemination and community mobilisation around social and political

[45] Lawrence Lessig, *Remix: Making Art and Commerce Thrive in the Hybrid Economy* (Penguin Press, New York, 2008).
[46] Henry Jenkins, *Textual Poachers: Television Fans and Participatory Culture* (Routledge, New York, 1992).
[47] Abigail De Kosnik, *Rogue Archives: Digital Cultural Memory and Media Fandom* (The MIT Press, Cambridge, Massachusetts, 2016).
[48] *Supra* note 42 at 12.
[49] Matthew Freeman, *Historicising Transmedia Storytelling: Early Twentieth-Century Transmedia Story Worlds* (Routledge, 2016).

issues.[50] IP enforcement may delete or restrict activist content under copyright claims, stifling these activities.

The Long Tail

Chris Anderson's Long Tail theory says that the internet has helped niche businesses and lesser-known content find an audience, changing the focus from mass-market hits to a wider range of cultural items.[51] Long Tail theory shows how digital platforms might help tiny creators reach an audience in an increasingly decentralised and democratised cultural landscape. IP laws may need to be changed to help producers and consumers in this wider cultural ecology due to niche content.[52]

Participatory culture represents a shift from traditional, top-down media consumption to a more interactive and democratised process, where users are creators, collaborators, and influencers. This cultural shift has been enabled by digital technologies, which provide the tools and platforms for widespread participation.

2. Looking into - Analysis of Findings

Algorithmic content moderation refers to the use of automated systems, often powered by machine learning algorithms, to review, filter, and manage content on digital platforms like social media, forums, and websites. The goal is to enforce community guidelines, legal requirements, and platform policies at scale.

2.1. Freedom of Expression - Facts, Ethical Considerations, Technology and Tools

Freedom of expression is essential to participatory culture, allowing individuals to share ideas and art without fear of censorship. However, AI-driven content moderation raises concerns about its impact on this freedom. While these systems can mistakenly censor legitimate content and contribute to echo chambers, they also have the potential to enhance freedom of expression by effectively filtering harmful content, supporting fact-checking, and providing privacy protection.[53]

---

[50] W. Lance Bennett and Alexandra Segerberg, *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics* (Cambridge University Press, 2012).

[51] Robert Armstrong, "The Long Tail: Why the Future of Business Is Selling Less of More", 33 *Canadian Journal of Communication* 127 (2008).

[52] *Ibid*.

[53] Barrie Sander, "Freedom of Expression in the Age of Online Platforms: The Promise and Pitfalls of a Human Rights-Based Approach to Content Moderation", 43 *Fordham International Law Journal* 939 (2020).

2.1.1.Facts and Challenges

Algorithmic content moderation systems are designed to identify and remove content that violates IP rights. These systems use various techniques, including content fingerprinting, hash matching, and machine learning algorithms, to detect potential infringements. While these technologies offer efficiency and scalability, they also present challenges:

Accuracy and Bias

Although designed to efficiently identify and delete copyrighted content, these algorithms often lack precision. Often, legitimate content is falsely tagged as infringing. This reduces system accuracy and reinforces biases, especially when algorithms are trained on data that may represent societal prejudices. Inaccuracies can stifle voices and target underrepresented communities, creating ethical questions about algorithmic enforcement's fairness and inclusivity.

Over-Moderation v. Under-Moderation

Algorithmic systems must balance over- and under-moderation. Over-moderation occurs when algorithms are too strict, removing content that does not breach IP rights. This can limit creativity and free expression, especially in transformational works like satire or critique. However, under-moderation occurs when algorithms fail to recognise infringing content, allowing it to spread. This damages rights holders and lowers platform integrity. Since over- and under-moderation can have serious repercussions, finding the appropriate balance is vital yet challenging.

Transparency and Accountability

Algorithmic content moderation lacks transparency and accountability. Users and authors rarely receive clear explanations for material removal or restriction. This opacity can damage moderation system confidence and cause injustice. Users have little recourse if their content was filtered incorrectly due to the lack of accountability procedures. The lack of clear guidelines and accessible appeal mechanisms makes many users feel helpless and alienated.

Context Sensitivity

Current algorithmic moderation systems struggle to recognise context, which is a major drawback. Satire, parody, and remix culture often depend on context. However,

algorithms use patterns and keywords without the complexity needed to discern content intent. Lack of context sensitivity can lead to erroneous enforcement measures, such as removing fair use or transformative use content.

Evolving Content Types

The ever-changing nature of digital content makes algorithmic filtering difficult. As memes, GIFs, deepfakes, and other multimedia formats arise, algorithms may struggle to keep up. Algorithms struggle to moderate these changing material kinds, either under- or over-moderated. Moderation systems must be updated and refined in this dynamic context, which is resource-intensive and technologically difficult.

These challenges highlight the need for a more nuanced and flexible approach to algorithmic content moderation, one that balances the protection of IP rights with the preservation of cultural expression and fairness in digital platforms.

2.1.2. Ethical Considerations

This subsection addresses the key ethical issues in algorithmic content moderation that explores the balance between protecting freedom of expression and preventing harm, the risks of biased or unjust content removal, and the challenges of integrating human oversight with AI systems.

Free Speech

Free speech is a major ethical issue in algorithmic content management. As digital platforms use algorithms to monitor and filter material, they may accidentally censor valid speech. By definition, algorithms follow preset rules and patterns, which may not adequately reflect human communication. This can unfairly remove free speech-protected content.[54] The moral challenge is reconciling the right to free speech with the necessity to delete dangerous content. Keeping algorithms from stifling open speech is a major challenge that requires serious attention and constant adjustment.

Censorship Concerns

Censorship concerns are linked to free speech. Uncalibrated algorithmic methods can restrict content that is contentious or politically sensitive but vital for public

[54] Jack M Balkin, "Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation", 51 *UC Davis Law Review* 1149 (2018).

debate. In circumstances where governments or powerful entities influence digital platforms, dissenting voices or minority perspectives may be suppressed. Content management should not violate the ability to access varied information and opinions. Preventing power abuses and protecting public discourse requires transparency in algorithm operation and content flagging and removal criteria.

Discrimination

Algorithmic content moderation discrimination is another major ethical issue. Large datasets may have biases reflecting social inequality while training algorithms. These mechanisms can perpetuate and worsen discrimination against marginalised populations.[55] Due to biased training data or a lack of cultural context, content from or about certain cultures may be flagged or removed. The ethical problem is to create and deploy fair and inclusive algorithms that treat all users equally. Diversity in training data, continuing review of algorithmic outcomes, and bias correction are needed.

Human-AI Collaboration or Hybrid Systems

Human-AI collaboration can improve moderation by combining algorithm efficiency with human contextual awareness. The possibility for exploitation or psychological harm when human moderators assess vast amounts of unpleasant content raises ethical concerns. Human monitoring poses accountability concerns, especially when human decisions overrule computer advice. Hybrid systems may cause complacency in monitoring their performance due to a false sense of security. Accountability and transparency require explicit human-machine responsibility distribution. Hybrid systems must be adaptable, responsive to input, and able to evolve with the digital information they manage for ethical reasons.[56]

2.1.3. Technology and Tools

From statistics and computer science, algorithmic content moderation employs a spectrum of approaches with varied complexity and efficacy. All of them seek to match, identify, forecast, or classify some piece of content (text, audio, image or video) based on its particular characteristics or general traits. Nevertheless, depending on the type of

[55] Ruha Benjamin, *Race after Technology: Abolitionist Tools for the New Jim Code Social Forces* (Polity Press, 2019).

[56] Bruno Lepri, Nuria Oliver, *et.al*., "Fair, Transparent, and Accountable Algorithmic Decision-making Processes", 31 *Philosophy & Technology* 612 (2018).

matching or classification needed, the approaches applied vary significantly as well as the kinds of data taken into account.[57] Various technologies and tools are employed in algorithmic content moderation to enforce IP rights:

Content ID Systems

Platforms like YouTube use Content ID systems to automatically manage copyrighted content. These systems generate digital fingerprints of copyrighted materials and compare them against a database of registered works. If a match is found, Content ID can automatically block, monetise, or track the content, helping to enforce copyright protection efficiently and with minimal manual oversight.

Hash Matching

Hash matching works by generating unique identifiers, or hashes, for digital content and comparing them to known infringing materials. This method helps platforms identify and block pirated content.[58] For example, Microsoft's PhotoDNA program breaks an image into a grid, resizes it to a standard size, converts it to black and white, and then computes a hash based on the intensity of each black-and-white section. This process allows PhotoDNA to effectively detect and manage unauthorised images.[59]

Machine Learning Algorithms

Machine learning algorithms analyse patterns in data to identify potential IP violations. These algorithms can be trained to recognize specific types of content, such as images, audio, and text, enhancing their ability to enforce IP rights.

Natural Language Processing (NLP)

For text-based IP enforcement, NLP systems are trained to evaluate human language and find writing violations. These algorithms compare text to vast copyright databases to find illegal copies of books, papers, and other materials.[60] They can discover

---

57 Robert Gorwa, Reuben Binns and Christian Katzenbach, "Algorithmic content moderation: Technical and political challenges in the automation of platform governance," 7(1) *Big Data and Society* (2020).

58 Kinza Yasar, "Hashing", *available at*: https://www.techtarget.com/searchdatamanagement/definition/hashing (last visited on February 05, 2025).

59 "PhotoDNA", *available at*: https://www.microsoft.com/en-us/photodna (last visited on February 05, 2025).

60 Fred Turner, "Burning Man at Google: A cultural infrastructure for new media production," 11 *New Media and Society* 73–94 (2009).

violations in paraphrased or altered information. Satire and parody, where meaning and context matter, may be difficult for NLP systems to understand, making IP enforcement problematic.

Image and Video Recognition

Visual media IP enforcement requires image and video recognition algorithms. These algorithms recognise copyrighted photographs, logos, and movies by assessing media visual patterns and features. YouTube flags or removes IP-infringing videos using content ID methods.[61] Illicit picture, artwork, and other visual material use can be detected using image recognition algorithms. These methods are effective, but they struggle to detect fair use-protected modified or derivative works like fan art or transformative videos.

Real-Time Filtering

Real-time filtering improves IP enforcement automation. This system immediately analyses uploaded content to avoid infringement. Real-time filtering algorithms block, flag, or restrict access to new content based on copyrighted works databases.[62] This proactive IP enforcement technique decreases illegal content but risks overkill and false positives by blocking lawful content.

User Reporting Integration

Content control relies on user reporting, although algorithms enforce IP. User reporting complements algorithmic systems by flagging IP-infringing content. User reporting and automated systems mix algorithm efficiency with human judgment. Collaboration can improve IP enforcement, but it needs good tools to handle and analyse user reports to prevent abuse or malicious reporting.

Algorithmic content moderation is essential for managing the vast amounts of content generated on digital platforms. However, it comes with significant challenges related to accuracy, bias, transparency, and ethics. The future of content moderation likely involves a combination of advanced algorithms, human oversight, and possibly new regulatory frameworks to ensure that content moderation is both effective and fair.

---

[61] European Union Intellectual Property Office, "Automated Content Recognition: Discussion Paper – Phase 2 IP enforcement and management use cases", 6 (2022).

[62] Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* 181 (Yale University Press, 2018).

2.2. Impact of Algorithmic Content Moderation on IP Enforcement

Algorithmic content moderation plays a significant role in IP enforcement by automating the detection and removal of infringing content. This approach offers several advantages:

Efficiency and Scalability: Algorithms can process vast amounts of data in real-time, offering a level of efficiency and scalability that human moderation cannot achieve. This is particularly important for platforms with millions of users and daily uploads.

Cost-Effectiveness: Automated systems reduce the need for human moderators, lowering operational costs for platforms. This cost-effectiveness is appealing to companies seeking to balance enforcement with profitability.

Consistency: Algorithms provide consistent enforcement of IP policies, reducing the variability and subjectivity associated with human decision-making.

2.3. Cultural Impact Assessment through Case Studies

The *Lenz* v. *Universal Music Corp*.[63] in 2015 involved a video of young children dancing to a Prince song "Let's Go Crazy". The video was made by the mother Stephanie Lenz in her kitchen, which Universal Music tried to remove from YouTube under the DMCA. The issue was whether copyright holders must consider fair use before issuing takedown notices. The Ninth Circuit ruled that Universal Music had misrepresented the video by not evaluating fair use, establishing that such consideration is required.[64] This decision reinforced fair use protections and influenced later cases, such as *Andy Warhol Foundation for the Visual Arts, Inc.* v. *Goldsmith et.al.*[65] in 2023, which also addressed fair use.

YouTube's Content ID

YouTube's Content ID system is a key example of algorithmic IP enforcement, effectively managing millions of copyrighted works. However, it has faced criticism for false positives and over-blocking, particularly in cases of transformative works like parodies and educational content. YouTube's guidelines on fair use and copyright help

[63] *Lenz* v. *Universal Music Corp.*, 801 F.3d 1126 (9th Cir. 2015).
[64] "Lenz v. Universal Music Corp. - Ninth Circuit Requires Analysis of Fair Use Before Issuing of Takedown Notices", 129 *Harvard Law Review* 2289-2290 (2016).
[65] *Andy Warhol Foundation for the Visual Arts, Inc.* v. *Goldsmith et al.*, 598 U.S. 598 (2023)

creators ensure compliance, with Content ID scanning videos for copyrighted material. Rights holders can choose to monetise or remove flagged content, while creators can dispute claims by asserting fair use, arguing their content's transformative nature.[66]

Fan Fiction and Copyright Claims

Fan fiction, a popular participatory culture, struggles with algorithmic IP enforcement. Transformative fan fiction uses copyrighted characters and settings. Streaming platforms like *Wattpad* and *Archive of Our Own* offer fan fiction. *Netflix* and *Wattpad* are producing original shows based on popular fan fiction.[67] Automated systems may take down fan-created information as copyright violation, stifling creativity and fan community sharing.

Music Mashups and Sampling

Music remixes and video montages combine copyrighted materials to create new works. Mashups transform copyrighted pieces into fresh ones. Spotify, YouTube, and others offer mashups. The YouTube channel "DJ Earworm" mashes up the top 100 tunes each year. The music industry has had mixed results with algorithmic enforcement. Algorithms can detect illicit copyrighted music use, but they may restrict transformative compositions that involve sampling and remixing.

Google/Jigsaw's Perspective API

Perspective is a popular NLP method for assessing remark or post "toxicity". This open-source toolbox lets website operators, researchers, and others apply machine learning models to evaluate user comments for harm or offense. Perspective in comment moderation systems alerts users if their remarks may be poisonous, allowing them time to edit. Perspective has drawbacks. Soon after its debut, researchers uncovered ways to deceive the technology, letting nasty remarks pass. Studies have demonstrated that the instrument can be biased, misclassifying some racial groups disproportionately. The

[66] Maayan Perel and Niva Elkin-Koren, "Accountability in Algorithmic Copyright Enforcement", 19 *Stanford Technology Law Review* 510 (2016).

[67] Anamika Unni, "Fan Fiction" and Copyright Liabilities: An Analysis from the American Legal Perspective", *available at*: https://csipr.nliu.ac.in/miscellaneous/fan-fiction-and-copyright-liabilities-an-analysis-from-the-american-legal-perspective/ (last visited on February 20, 2025).

Conversation AI team behind Perspective advises against entirely automated moderation using the API due to its mistake rate.[68]

Despite these protections, creators still face challenges, as platforms' algorithms can mistakenly remove legitimate fair use content. Aggressive copyright enforcement by large media companies often suppresses transformative content, potentially stifling creativity and adding to legal uncertainties for content creators. Algorithmic IP enforcement has significant cultural implications, affecting creativity, innovation, and the sharing of cultural resources

2.4. Stakeholder Perspectives

This subsection explores the views of rights holders on the advantages of algorithmic enforcement for protecting their IP and deterring piracy. It also addresses the challenges faced by creators and users, including issues with false positives, complex IP laws, and the need for fair use considerations in algorithmic systems.

2.4.1. Rights Holders' Views on Algorithmic Enforcement

Rights holders, including artists, publishers, and media companies, have a vested interest in protecting their Intellectual Property. From their perspective, algorithmic content moderation offers several benefits:

Protection of Economic Interests

Algorithmic enforcement helps rights holders protect their economic interests by preventing unauthorised use of their works. Automated systems can identify and remove infringing content, safeguarding revenue streams.

Deterrence of Piracy

Algorithms can deter piracy by swiftly identifying and removing unauthorised copies of copyrighted content. This deterrence is crucial in maintaining the value of creative works in the digital marketplace.

[68] Emma Llansó, Joris van Hoboken, *et.al.*, "Artificial Intelligence, Content Moderation, and Freedom of Expression" 5 *(Transatlantic Working Group*, 2020).

Efficiency and Speed

Rights holders appreciate the efficiency and speed of algorithmic enforcement, which allows for quick responses to potential IP violations. This timeliness is vital in preventing the widespread distribution of infringing content.

2.4.2. Creators' and Users' Experiences and Challenges

Creators and users of participatory culture face unique challenges in navigating algorithmic IP enforcement:[69]

Unintended Consequences

Creators often experience unintended consequences of algorithmic enforcement, such as false positives and over-blocking. These errors can result in the removal of legitimate content and hinder creative expression.

Navigating Complexities

The complexities of IP laws and algorithmic systems can be daunting for creators, particularly those without legal expertise. The fear of enforcement may deter creators from engaging in transformative and remix culture.

Advocacy for Fair Use

Creators advocate for the recognition of fair use and transformative works within algorithmic enforcement systems. They argue that algorithms should be designed to account for the nuances of cultural expression and artistic transformation.

2.5. Moderation – Legitimacy, Surveillance and Errors

Critics have attacked social media platforms for their content screening as online and offline communication combine.[70] Platform algorithm design and content removal affect users, public opinion, and lawmakers. Public acceptance and moderator trust are needed by organisations and social media. If human or algorithmic content screening is unreliable, users and politicians may use platforms differently and break platform rules. Algorithmic content moderation raises authenticity, espionage, and IP enforcement

[69] Kristen Vaccaro, Christian Sandvig and Karrie Karahalios, "At the End of the Day Facebook Does What It Wants': How Users Experience Contesting Algorithmic Content Moderation", 4 *Proceedings of the ACM on Human Computer Interaction* 167:1–167:22 (2020).

[70] Jelizaveta Juřičková, "Enhancing Legitimacy of Content Moderation," 15 *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 1 (2024).

concerns. When IP rules trump user rights, algorithmic enforcement is questionable. Without transparency, algorithmic decision-making can weaken police efforts. Algorithmic systems' massive data collection and processing raise privacy and surveillance concerns. Monitoring and controlling user-generated content with algorithms may violate privacy and autonomy. Due to bias and errors, algorithmic systems can make unfair enforcement decisions. Due to algorithmic bias, moderation must be transparent and accountable.

2.6. Current Regulatory Efforts

Digital Millennium Copyright Act (DMCA)

The DMCA requires platforms to implement copyright enforcement tools to address internet IP infringement. DMCA's compliance focus has been criticised for fostering excessive content banning and hampering creativity. Its safe harbor protections protect online service providers from liability for user-generated content if they follow the notice-and-takedown process.[71] Critics say this method causes platforms to over-block content to avoid legal risks, stifling creative innovation and free expression.

EU Directive on Copyright in the Digital Single Market

The EU Directive aims to harmonise copyright rules and strengthen online IP enforcement. It contains openness and accountability principles for platforms. Critics say it could damage user rights and free speech. Critics fear that the necessity for proactive steps may lead to widespread usage of content recognition technologies, which could limit user-generated material and violate privacy.[72] The Directive's adoption has sparked discussions on how to safeguard copyright holders while preserving a vibrant digital culture that encourages remixing and creative expression.[73]

3. Looking Ahead - Discussions and Implications

This section examines the future of balancing IP enforcement with cultural expression. It highlights the need for improved algorithmic accuracy, transparency, and

[71] Martin Senftleben, "Bermuda Triangle: Licensing, Filtering and Privileging User Generated Content Under the New Directive on Copyright in the Digital Single Market," 35 *Computer Law & Security Review* 2 (2019).

[72] Julia Reda, "The EU Copyright Directive: A Balanced Approach or a Missed Opportunity?" 14 *Journal of Intellectual Property Law & Practice* 813–819 (2019).

[73] João Pedro Quintais, "The New Copyright in the Digital Single Market Directive: A Critical Look," 42 *European Intellectual Property Review* 28–41 (2020).

stakeholder collaboration. Recommendations are provided for policymakers and platforms to enhance regulation, refine AI systems, and empower users.

3.1. Navigating the Balance; IP Enforcement and Cultural Expression

Integrating IP enforcement and participatory culture is tough. Commercial interests, artistic expression, and cultural exchange must be balanced in IP enforcement. Fair use and transformative works must be recognised since strict enforcement hinders innovation and cultural sharing. Consider fair usage while using algorithms to avoid eliminating legitimate content. Trust in automated judgments demands transparency and accountability, and platforms must offer clear appeals for incorrect removals. Current algorithms require human review for fairness due to false positives and context concerns. Creators and rights holders can assist create more inclusive and effective enforcement mechanisms.

3.2. Implications for Policy and Practice

This sub-section outlines key recommendations for policymakers and platforms to address the challenges of algorithmic content moderation and IP enforcement.

3.2.1. Recommendations for Policymakers

Regulation

We need strong regulations to balance creator, platform, and user rights as AI is used in content filtering and IP enforcement. Government control of algorithmic systems should be researched to ensure openness, accountability, and justice. Guidelines for creating, testing, and auditing algorithms and digital platforms' global, cross-border character are included. Effective AI legislation must protect IP, free speech, and information access while assuring ethical use.

Promote Fair Use and Transformative Works

Fair use and transformative works are essential to creativity, innovation, and cultural exchange. Policymakers should understand this. Creative expression, remix culture, and new storytelling are encouraged by fair usage. Transformative works foster a participatory culture where people can rewrite cultural narratives. The intricacies of new digital settings require revised legal frameworks to cater for cultural expression and artistic development.

Enhance Transparency and Accountability

Policymakers should advocate for transparency and accountability in algorithmic enforcement processes. Platforms should be required to provide clear explanations of enforcement decisions and mechanisms for appealing erroneous removals.

Encourage Stakeholder Collaboration:

Policymakers should facilitate collaboration among stakeholders to address the challenges of algorithmic enforcement. Collaborative approaches that involve rights holders, creators, and users can enhance the legitimacy and fairness of enforcement efforts.

3.2.2. Suggestions for Platforms

Improve Algorithmic Accuracy

For fair use and transformative works, platforms should increase algorithmic accuracy to avoid false positives and over-blocking. IP enforcement requires more complicated algorithms as AI advances without limiting inventiveness. AI research in NLP, picture, and video identification must improve context and intent understanding to distinguish parody and satire from infringement. Reduce algorithmic prejudice to treat cultural expressions fairly. These improvements can make AI a more reliable IP enforcement tool with less human intervention.

Implement Human Oversight

In order to overcome the limits of algorithmic systems, platforms ought to enact procedures for human monitoring. It is possible for human moderators to provide contextual awareness and judgment when making decisions regarding enforcement.

Enhance User Empowerment

Platforms should disclose legal intricacies to explain IP rights and enforcement. Artists can educate users to manage algorithmic enforcement. A fair digital ecology requires user empowerment. Explore user content moderation to empower users with content control and platform algorithm engagement. Developers can appeal moderation, adjust content preferences, and improve algorithms *via* feedback. Research should educate customers about AI, algorithmic decision-making, and digital rights. Empowered

users can educate and communicate with the community, minimising over- and under-moderation and digital platform trust.

4. The Way Forward

This piece examines the complex interplay between algorithmic content regulation, IP enforcement, and digital participatory culture. The study highlights a contradiction between preserving rights holders' economic interests and promoting creative expression and cultural exchange in an AI-driven content regulatory world. This careful balance ensures that creativity is not strangled and that digital platforms remain areas where users can actively contribute to cultural creation without restrictive enforcement methods.

In the research, algorithmic bias and prejudice are major concerns. Despite their speed and scalability, automated systems suffer with IP law, especially fair use and transformative works. When trained on huge datasets, algorithms can accidentally include biases that disproportionately affect specific tribes or cultures. This can unfairly target or underrepresent certain populations, presenting ethical and legal concerns. These algorithms may misclassify transformative content like fan fiction, remixes, and memes as copyright infringement, limiting artistic freedom and participatory culture, according to the study.

The research analyses real-world instances like *Lenz* v. *Universal Music Corp.* to show how fair usage is complicated and how automated enforcement techniques typically fail. The paper also examines the Supreme Court's *Andy Warhol Foundation* v. *Goldsmith* decision, highlighting the difficulties of differentiating infringement from transformative expression. These instances and stakeholder involvement from rights holders, producers, and users demonstrate algorithmic IP enforcement's broad issues. The investigation found various technology and policy shortcomings. Current algorithmic algorithms cannot interpret context, meaning, and intent, which are essential in fair use scenarios. This leads to false positives and over-blocking, which stifles innovative content categories. These algorithms are opaque, so users and artists have no idea why their content is flagged or removed, which lowers trust in these platforms and weakens their impartiality, according to the research.

The paper proposes policy changes to make algorithmic enforcement more fair. It requires strong regulatory frameworks to assure transparency, accountability, and justice in system deployment. Fair use and transformative works are crucial to creativity and innovation, so policymakers should recognise them. Rights holders, producers, and platforms should work together to create effective, just, and culturally appropriate enforcement procedures, according to the report. The research also stresses the importance of human moderation oversight. AI technologies promise remarkable efficiency, but they cannot yet match human moderators' nuanced judgment, especially in complex IP enforcement. Human monitoring can improve algorithm performance and ensure contextually sound and equitable conclusions.

Future research should improve AI technologies to understand cultural context, reduce algorithmic bias, and ensure that IP enforcement mechanisms do not disproportionately affect marginalised communities or inhibit creative expression. Since digital platforms and content moderation are global, the report recommends more investigation into their international regulation. This research shows that algorithmic IP enforcement must combine creators' rights with participatory culture's creativity and expression. AI will continue to shape the digital landscape, making it crucial to create smart, ethical frameworks that promote innovation and fairness. Addressing technical, policy, and practice gaps can help stakeholders create a digital ecosystem where creativity, culture, and commerce coexist.