

Balancing scale and fairness: The societal impact of AI-powered enforcement systems

Prabhakar Singh *

Meta, USA.

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(02), 2610–2617

Publication history: Received on 04 April 2025; revised on 20 May 2025; accepted on 22 May 2025

Article DOI: <https://doi.org/10.30574/wjaets.2025.15.2.0817>

Abstract

This article explores the societal implications of AI-powered enforcement systems that govern user behavior on digital platforms such as marketplaces, social networks, and gig economy services. As these platforms have scaled to serve billions of users globally, they have transitioned from manual content moderation to sophisticated automated systems that make enforcement decisions at unprecedented scale. While these systems effectively detect policy violations, they create complex tensions between platform protection and user fairness. The article examines how enforcement actions impact users' economic outcomes, trust perceptions, and equitable access to digital opportunities, with particular attention to potential disparate impacts on underrepresented communities. It presents the fundamental dilemma platforms face between robust protection and precision, analyzes emerging approaches to more responsible enforcement, and advocates for enforcement systems designed with both technical and ethical considerations. By adopting contextual enforcement, meaningful transparency, accessible appeals, and inclusive design principles, platforms can develop governance mechanisms that maintain ecosystem integrity while ensuring fair treatment for all participants.

Keywords: Algorithmic fairness; Content moderation; Platform governance; Digital rights; Automated enforcement

1. Introduction

In today's digital landscape, platforms serving billions of users face an unprecedented challenge: maintaining trust and safety at scale while ensuring fairness for all participants. As marketplaces, social networks, and gig economy platforms have grown into global infrastructure, the systems that govern user behavior have evolved from simple manual reviews to sophisticated AI-powered enforcement mechanisms that make millions of decisions per day.

The scale of these operations is staggering. Major social media platforms' Community Standards Enforcement Reports reveal that their systems evaluate content across dozens of languages and regional contexts, making billions of enforcement decisions annually that directly impact user experiences and participation across their various services ^[1]. Beyond social media, the OECD notes that e-commerce platforms deploy sophisticated algorithmic systems to screen listings, monitor seller behavior, and identify potentially fraudulent activities among millions of daily transactions, leveraging both machine learning and rule-based approaches to maintain marketplace integrity ^[2].

While technically impressive, these systems raise important questions about fairness and equity. Meta's transparency reporting has begun addressing these concerns by including appeal rates and restoration metrics, acknowledging the importance of measuring false positives and false negatives in enforcement actions ^[1]. Meanwhile, the OECD's analysis highlights how small businesses and independent service providers may be disproportionately affected by false enforcement actions, often lacking the resources to effectively navigate appeal processes compared to larger market participants ^[2].

* Corresponding author: Prabhakar Singh

As platforms continue to refine these systems, they are increasingly adopting more nuanced approaches. Meta has documented their shift toward "graduated enforcement" frameworks that apply proportionate responses based on violation severity, user history, and content context [1]. Similarly, the OECD describes how leading digital marketplaces have implemented tiered enforcement protocols that escalate gradually from warnings to restrictions to account termination, allowing sellers opportunities to address issues before facing severe consequences [2].

This evolution reflects a growing recognition that effective platform governance requires both technological sophistication and careful consideration of its broader economic and social impacts. The most successful platforms will be those that can maintain ecosystem integrity while ensuring that enforcement systems treat all participants fairly and equitably.

2. The Rise of Automated Enforcement

The sheer volume of interactions on major platforms has made human review of potential violations impractical. Digital platform operations now occur at a truly unprecedented scale. According to one popular photo-sharing platform's engineering team, their systems process over 95 million photos and videos daily, while a leading video-sharing site reports that users upload more than 500 hours of video content every minute. This massive volume represents just a fraction of the content that major platforms must evaluate for policy compliance. Research on platform governance and content moderation indicates that major social media platforms must evaluate billions of pieces of content weekly, with automated systems now handling the vast majority of initial screening [3]. These systems have evolved rapidly as platforms face mounting pressure to address harmful content while managing astronomical content volumes that make comprehensive human review logistically impossible. The economic realities of content moderation at scale have pushed platforms toward increasingly sophisticated automated approaches, as maintaining teams of human moderators large enough to manually review all potentially problematic content would be prohibitively expensive. The computational demands extend beyond social media into commerce and service platforms, where transactions and interactions must be continually evaluated for potential policy violations, creating technical challenges that only machine learning approaches can feasibly address.

Table 1 Technical Evolution of Platform Enforcement Systems ^{3,4}

Platform Type	Content Volume	Primary Technical Approaches	Key Challenges
Social Media	Billions of daily uploads	Machine learning classifiers, Natural language processing	Cross-cultural context challenges
E-commerce	Millions of hourly listings	Risk scoring, Pattern recognition	Balancing fraud prevention with seller experience
Gig Economy	Continuous transaction validation	Graph-based algorithms, Behavioral analysis	Ensuring contextual understanding of transactions
Content Sharing	Hundreds of hours per minute	Multi-layered enforcement, Human-AI collaboration	Maintaining user agency when errors occur
Messaging	Private communications at scale	Signal detection, Metadata analysis	Preserving privacy while detecting harmful content

This scale necessitates AI-powered enforcement systems that can detect patterns of spam, fraud, and abuse in real-time. The technical approaches have evolved significantly in recent years, moving beyond simple rule-based detection. Modern enforcement systems leverage supervised machine learning classifiers trained on vast datasets of previously identified violations, allowing them to recognize subtle patterns that might indicate problematic content or behavior. These systems increasingly integrate natural language processing capabilities to understand nuanced content meaning across multiple languages, enabling more precise policy enforcement across diverse global communities. The Tech Policy Press's analysis of content moderation systems emphasizes that platforms have increasingly developed multi-layered enforcement approaches that combine algorithmic detection with varying levels of human review based on risk assessment. Their research highlights how platforms have invested in developing increasingly context-aware detection systems that consider not just the content itself but surrounding signals including user history, interaction patterns, and community-specific norms. Despite these advances, they note significant challenges remain in striking the appropriate balance between automated enforcement efficiency and ensuring users maintain appropriate agency and appeal rights

when enforcement errors occur. The research particularly emphasizes how automated systems face challenges with contextual understanding across different cultural settings, as platforms struggle to develop the scaling capacity needed to fully comprehend the nuances of local languages, humor, traditions, and political cultures that shape users' perceptions of online speech globally ^[4].

3. Beyond Technical Performance: The Human Impact

While engineering teams typically optimize for precision metrics like false positive rates and detection accuracy, the real-world impact of these systems extends far beyond technical benchmarks. Enforcement actions directly affect users in profound ways that are often overlooked in purely technical evaluations of system performance. Research examining the human dimensions of content moderation reveals the significant psychological and economic toll these systems can take on users caught in enforcement actions. As Steve Carey notes in his analysis of content moderation systems, "Behind every flagged post, suspended account, or removed listing is a human being whose digital presence—and often livelihood—hangs in the balance." His interviews with platform users who experienced enforcement actions highlight that even temporary restrictions can leave lasting impacts on users' financial stability and emotional well-being. This perspective represents an important shift in how we evaluate platform governance, moving beyond technical metrics toward frameworks that consider the full spectrum of human consequences ^[5].

3.1. Economic Outcomes

When marketplace sellers face account restrictions or payment holds, their livelihoods can be disrupted. Small businesses that depend on platform visibility may experience devastating financial consequences from algorithmic decisions. For many digital entrepreneurs, particularly those in emerging economies, platforms represent their primary or sole income source, making enforcement actions particularly consequential. Steve Carey's research documented cases where algorithm-driven account restrictions led to severe financial hardship, with some affected sellers reporting loss of their entire income stream during appeal periods lasting weeks or even months. The economic impact is particularly severe for sellers with limited financial reserves, creating cascading effects that extend beyond the immediate lost revenue to affect their ability to maintain supplier relationships, cover business expenses, and support dependents. This highlights the significant real-world consequences of false positives in enforcement systems, where technical "errors" translate directly into economic instability for real people ^[5].

3.2. Trust and Perception

Users subjected to false enforcement actions often develop lasting negative perceptions about platform fairness. This "enforcement trauma" can persist even after successful appeals, leading to diminished trust in the platform's governance. Steve Carey's analysis reveals that users who experience enforcement actions they believe were wrongful demonstrate significantly reduced platform engagement even after their accounts are restored. His interviews with affected users documented emotional responses ranging from frustration to helplessness, particularly when navigating opaque appeal processes with limited human interaction. The impact of these experiences extends beyond individual users to affect broader community perception, as users share their negative experiences through social networks and professional communities, creating ripple effects that can damage the platform's reputation among entire user segments. This trust deficit carries long-term consequences for both individual platform participation and broader ecosystem health ^[5].

3.3. Access and Inclusion

Enforcement systems that fail to account for cultural, linguistic, or contextual nuances may disproportionately impact underrepresented groups. What appears as anomalous behavior might simply reflect different usage patterns among diverse communities. Recent legal developments highlight growing awareness of these disparate impact concerns. The Consumer Financial Protection Bureau has emphasized that automated decision systems, including those used for marketplace governance, can potentially create unintended barriers for certain demographic groups even without explicit discriminatory intent. As discussed in legal analyses of the regulatory landscape, algorithmic systems trained primarily on data from dominant user populations may inadvertently encode biases that result in higher enforcement rates for underrepresented communities. These concerns have led to increased regulatory scrutiny of automated decision systems, with both government agencies and platform operators exploring approaches to measuring and mitigating disparate impact. Such efforts reflect recognition that ensuring equitable access to digital economic opportunities requires careful attention to how enforcement systems affect different user populations ^[6].

Table 2 Human Consequences of Automated Enforcement Systems ^{5,6}

Impact Dimension	Primary Affected Users	Manifestation	Long-term Consequences
Economic Outcomes	Small businesses, Independent sellers	Lost revenue, Payment holds, Reduced visibility	Financial instability, Business closures, Market exit
Trust and Perception	Users experiencing enforcement	Frustration, Helplessness, Reduced engagement	Platform abandonment, Community skepticism, Reputation damage
Access and Inclusion	Underrepresented groups, Linguistic minorities	Higher enforcement rates, Cultural misunderstandings	Digital exclusion, Reinforced inequalities, Reduced diversity
Platform Health	Entire user base	Ecosystem quality, Safety perception	User retention, Community vitality, Platform sustainability
Regulatory Risk	Platform operators	Compliance challenges, Legal exposure	Increased oversight, Operational constraints, Market limitations

4. The Enforcement Dilemma: Protection vs. Precision

Platform teams face a fundamental tension between two competing objectives: protecting the ecosystem from genuine abuse that harms users and undermines platform integrity, while minimizing collateral damage to legitimate users caught in enforcement actions. This tension represents what governance researchers describe as the "moderation dilemma" - a complex balancing act with significant consequences in either direction. As explained by Maxicus in their comprehensive analysis of social media content moderation approaches, this challenge has intensified as platforms scale globally. Their research highlights how content moderation systems must simultaneously filter out harmful content while preserving authentic user expression. "Social media platforms must maintain a delicate balance between removing harmful content and allowing free expression," they note, observing that platforms typically operate along a spectrum of enforcement strictness, with different strategic approaches based on their risk tolerance and user base characteristics. Platforms serving vulnerable populations or handling sensitive content typically implement more aggressive enforcement systems, accepting higher false positive rates to minimize potential user harm, while platforms prioritizing open expression might take more measured approaches ^[7].

This creates difficult tradeoffs that platform governance teams must navigate daily. Strict enforcement reduces abuse but increases false positives. According to Gokan Ozcifci's analysis of scalable governance approaches, attempting to eliminate all policy violations inevitably leads to overenforcement that damages legitimate platform activity. Their research into governance frameworks emphasizes that effective enforcement requires proportionality rather than zero-tolerance approaches. "As organizations scale, governance cannot be one-size-fits-all," they observe, noting that the most effective governance systems adapt enforcement stringency based on risk levels and potential impact. Their guide documents how platforms attempt to mitigate these tradeoffs through various technical and procedural approaches, including tiered enforcement mechanisms, confidence thresholds that trigger different levels of review, and specialized workflows for borderline cases where algorithmic decisions have lower confidence. This perspective reflects a growing recognition that enforcement is not a binary decision but rather a spectrum of possible interventions proportionate to both violation severity and detection confidence ^[8].

Lenient systems may preserve user experience but allow more abusive content or transactions to slip through. Maxicus's analysis quantifies this relationship, noting that platforms with more permissive moderation typically experience higher rates of user reports about harmful content, creating different but equally significant trust challenges. Their research documents how visible policy violations can drive user dissatisfaction and platform abandonment, particularly among vulnerable user groups who may be disproportionately affected by harmful content. "When platforms fail to adequately enforce their policies, they risk creating hostile environments that exclude certain communities," they observe, highlighting the reputational and ethical risks of underenforcement. Maxicus's work similarly emphasizes that governance systems must find appropriate balance points rather than maximizing for a single metric. Their guide suggests that platforms increasingly recognize this is not a zero-sum tradeoff but rather requires sophisticated governance approaches that can adapt enforcement stringency based on violation type, potential harm, and community context, recommending multi-layered approaches that balance automation with human judgment ^[7].

Table 3 Strategic Tradeoffs in Platform Enforcement Approaches ^{7,8}

Enforcement Approach	Ecosystem Protection	User Experience	Implementation Mechanisms	Platform Examples
Zero Tolerance	Maximum protection, Minimal policy violations	High false positives, Restrictive environment	Strict automated filters, Pre-publication review	Financial services, Child safety platforms
Balanced Enforcement	Moderate protection, Some violations detected	Occasional false positives, Reasonable restrictions	Tiered enforcement, Confidence thresholds, Mixed review	E-commerce, Professional networks
Community-Led	Basic protection, Focus on egregious violations	Minimal false positives, Open expression	User reporting, Community standards, Post-publication review	Discussion forums, Creative communities
Contextual	Targeted protection, Nuanced violation detection	Reduced false positives, Personalized experience	Machine learning with context awareness, Reputation systems	Leading social platforms, Modern marketplaces
Risk-Based	Variable protection based on assessment	Differentiated experience by risk level	Real-time scoring, Progressive interventions	Emerging hybrid approaches

5. Building More Responsible Enforcement Systems

Forward-thinking platforms are adopting several approaches to balance these competing priorities, recognizing that responsible enforcement requires both technical sophistication and thoughtful governance frameworks. The evolution toward more nuanced enforcement systems reflects growing awareness of both the power and limitations of automated moderation tools, as well as increasing expectations from users, regulators, and civil society organizations for more equitable platform governance. As platforms scale to serve global audiences, the need for enforcement approaches that can accommodate diverse contexts while maintaining consistent policy application has become increasingly apparent to industry leaders and governance researchers alike ^[9].

5.1. Contextual Enforcement

Modern enforcement systems must move beyond one-size-fits-all rules to consider multiple factors that influence the appropriate response to potential violations. Iberdrola's analysis of digital rights in modern platforms emphasizes that contextual enforcement represents a critical evolution in how platforms balance safety and user autonomy. Their research notes that "digital rights in the context of online platforms include the right to fair and transparent treatment by automated systems," highlighting how platforms increasingly incorporate user history and established reputation into enforcement decisions. This approach enables systems to distinguish between isolated incidents and pattern-based abuse, allowing for more proportionate responses. Research on platform governance similarly underscores the importance of regional and cultural context in enforcement decisions, noting that "content moderation systems must account for linguistic and cultural diversity to ensure equitable treatment for users across all regions and communities." This analysis documents how advanced platforms now adjust enforcement thresholds based on the type of activity and potential harm, enabling more nuanced responses that match enforcement severity to violation impact rather than applying uniform penalties regardless of context ^[3, 2].

5.2. Transparency and Explanation

When enforcement actions occur, platforms should provide clear information to affected users about the basis for the decision. Iberdrola's digital rights framework identifies transparency as a foundational principle for responsible platform governance, stating that "users have the right to understand decisions that affect their digital participation." Their analysis documents how leading platforms have evolved from generic policy notifications toward more specific explanations that help users understand precisely what triggered enforcement actions. Research on content moderation similarly highlights the importance of specific policy references rather than generic notifications, noting that "transparent communication about moderation decisions builds user trust and reduces appeal volumes." Both research

teams emphasize that effective transparency involves striking a delicate balance—providing sufficient detail for understanding while avoiding technical language that might confuse non-expert users or reveal system vulnerabilities that could enable policy circumvention [3, 2].

5.3. Meaningful Appeal Mechanisms

Even the best systems make mistakes. Effective appeal processes require thoughtful design to ensure they provide genuine recourse rather than merely procedural compliance. Iberdrola's analysis of digital rights emphasizes that "the right to contest automated decisions is fundamental to fair platform governance," documenting how accessible appeal mechanisms serve as a critical accountability check on automated enforcement systems. Their research highlights how leading platforms have implemented multiple review paths for appeals, combining automated reassessment for straightforward cases with human review for complex situations. Saad Khan and Mia Lucas's examination of content moderation systems similarly underscores the importance of timeliness in appeals processing, noting that "appeal mechanisms that take weeks to resolve can cause irreparable harm to users whose livelihoods depend on platform access." Both analyses emphasize that effective appeals processes must include clear provisions for reputation restoration when errors are identified, ensuring that past enforcement mistakes don't continue to influence future algorithmic decisions [3, 2].

5.4. Inclusive Design Principles

Reducing algorithmic bias requires deliberate design choices throughout the development and implementation of enforcement systems. Iberdrola's digital rights framework identifies non-discrimination as a core principle, stating that "automated systems must not perpetuate or amplify existing social biases." Their analysis emphasizes the importance of diverse training data representing global user populations, noting that systems trained primarily on content from dominant groups tend to produce higher error rates when evaluating content from underrepresented communities. Saad Khan and Mia Lucas's research on platform governance similarly highlights the value of regular bias audits across demographic groups, documenting how "proactive monitoring for disparate impact can identify and address enforcement patterns that disproportionately affect specific user communities." Both research efforts emphasize the importance of cross-functional teams that include social scientists and ethicists alongside technical experts, bringing diverse perspectives to the enforcement system design that can anticipate potential equity concerns throughout the development process [3, 2].

Table 4 Framework Components for Responsible Enforcement Systems 3.2

Design Principle	Key Components	Implementation Strategies	Expected Benefits
Contextual Enforcement	User history, Regional context, Violation type, Proportional response	Reputation systems, Cultural calibration, Harm assessment, Graduated actions	Reduced false positives, More appropriate interventions, Increased perceived fairness
Transparency & Explanation	Clear communication, Specific policy references, Accessible language, Appropriate detail	Detailed notifications, Policy education, Simplified explanations, Multi-format communications	Improved user understanding, Reduced appeals, Increased compliance, Trust building
Meaningful Appeals	Low-friction options, Multiple review paths, Timely resolution, Reputation restoration	One-click appeals, Automated + human review, Priority systems, Record correction	Error correction, Reduced economic impact, User confidence, System improvement
Inclusive Design	Diverse training data, Regular bias audits, Community engagement, Cross-functional teams	Global data collection, Demographic analysis, Stakeholder consultation, Interdisciplinary development	Equitable outcomes, Reduced disparities, Broader accessibility, Better representation

6. The Path Forward

As AI enforcement systems become more sophisticated, their design must be informed not just by technical capabilities but by their broader societal impact. This perspective represents a significant evolution in how platforms approach governance, moving beyond purely technical optimization toward more holistic frameworks that consider enforcement

systems' effects on user communities, economic opportunity, and social equity. Research on responsible AI governance emphasizes that effective governance frameworks must address both the technical functionality and social dimensions of automated systems. As noted in Science Direct's analysis, "Responsible AI governance has been proposed as a comprehensive approach to manage AI development and deployment in ways that ensure AI systems' trustworthiness and safety." Their research documents how leading platforms have begun integrating social impact considerations into their enforcement system architecture, rather than treating them as separate concerns addressed only after technical development. This approach recognizes that enforcement systems simultaneously serve technical, social, and economic functions, necessitating design processes that consider these dimensions in tandem rather than sequentially ^[10].

This integrated approach requires interdisciplinary teams that bring together security engineers, data scientists, policy experts, and user advocates throughout the development process. Deepesh PC's analysis of platform governance best practices documents how organizations with diverse governance teams tend to build more balanced enforcement systems that better anticipate potential unintended consequences. Their research notes that "proper governance is not just about controlling access or implementing security measures; it's about creating a framework that enables innovation while ensuring compliance and managing risks." By integrating policy expertise, user advocacy, and ethical perspectives alongside technical capabilities, platforms typically develop enforcement architectures that more effectively balance protection and fairness. The Science Direct research similarly emphasizes how interdisciplinary approaches help platforms anticipate potential disparate impacts earlier in the development process, enabling more proactive mitigation strategies before systems are deployed at scale ^[10, 11].

Responsible enforcement also requires governance frameworks that balance innovation with responsibility. Deepesh PC's research highlights how leading organizations have established internal governance structures specifically designed to evaluate enforcement systems across multiple dimensions, noting that "governance in the digital transformation era is not about restricting possibilities but about channeling innovation in responsible directions." These governance frameworks typically include clear escalation paths for high-impact enforcement decisions, regular review processes to evaluate system performance across user segments, and accountability mechanisms that connect enforcement outcomes to broader platform objectives. Their analysis documents how such frameworks help platforms move beyond viewing enforcement solely as a technical function toward understanding it as a core governance responsibility with significant implications for platform health and user trust ^[11].

Ongoing dialogue with affected communities represents another critical component of responsible enforcement design. The Science Direct research emphasizes that engagement with stakeholders is essential for developing AI systems that align with user needs and societal values. Their analysis documents how platforms that maintain regular engagement with diverse user communities, particularly those who have experienced enforcement actions or represent frequently affected groups, tend to identify potential issues earlier and develop more effective mitigation strategies. This engagement can take multiple forms, from formal feedback mechanisms and policy consultations to ongoing dialogue with community representatives and advocacy organizations. When implemented effectively, these feedback loops create valuable sources of insight that technical metrics alone cannot provide, helping platforms understand the lived experience of enforcement from user perspectives ^[10].

Transparent reporting on enforcement outcomes serves as both an accountability mechanism and a driver of continuous improvement. Deepesh PC's analysis highlights how public transparency reporting has evolved from basic metrics toward more nuanced disclosures that provide insight into system performance and impact. Their research notes that "transparency creates a foundation of trust between platforms and their users, while also establishing accountability for governance outcomes." While appropriate transparency levels vary based on platform type and user base, the research emphasizes that reporting should provide meaningful insight into both the scale and impact of enforcement operations while respecting privacy considerations and avoiding disclosures that might enable system gaming ^[11].

By viewing enforcement systems through both technical and ethical lenses, platforms can build governance mechanisms that protect ecosystems while preserving the dignity and opportunity that digital platforms promise to all users. The most successful platforms will be those that recognize that trust, safety, and fairness are not competing priorities but interdependent values that must be designed into the core of automated enforcement systems. As the Science Direct research concludes, responsible governance approaches must "balance innovation with ethical considerations, ensuring that automated systems advance beneficial outcomes while minimizing potential harms." This balanced perspective represents the frontier of responsible enforcement design, pointing toward governance approaches that maintain platform integrity while ensuring equitable access to digital opportunity ^[10].

7. Conclusion

As AI enforcement systems continue to evolve in sophistication and reach, their design must transcend purely technical considerations to encompass broader societal impacts. The future of effective platform governance lies not in viewing trust, safety, and fairness as competing priorities but as interdependent values that must be integrated throughout enforcement system design. By assembling interdisciplinary teams, establishing balanced governance frameworks, maintaining ongoing dialogue with affected communities, and implementing transparent reporting mechanisms, platforms can build enforcement architectures that protect digital ecosystems while preserving user dignity and opportunity. The most successful platforms will be those that recognize enforcement as a sociotechnical system with profound implications for economic participation and digital inclusion, requiring approaches that advance both protection and fairness simultaneously. This balanced perspective represents the frontier of responsible enforcement design, pointing toward governance approaches that maintain platform integrity while ensuring equitable access to digital opportunities for all users across global communities.

References

- [1] Meta, "Community Standards Enforcement Report," Meta Transparency Center. [Online]. Available: <https://transparency.meta.com/reports/community-standards-enforcement/>
- [2] OECD, "The role of online marketplaces in enhancing consumer protection," 2021. [Online]. Available: https://goingdigital.oecd.org/data/notes/No7_ToolkitNote_ConsumerProtection.pdf
- [3] Saad Khan and Mia Lucas, "Platform Governance and Content Moderation: Examining the role of social media platforms in content moderation, including policies, guidelines, and challenges related to regulating news content," ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/372216340_Platform_Governance_and_Content_Moderation_Examining_the_role_of_social_media_platforms_in_content_moderation_including_policies_guidelines_and_challenges_related_to_regulating_news_content
- [4] Hiromitsu Higashi, "Right, Capacity and Will in Content Moderation: A Case for User Empowerment," Tech Policy Press, 2023. [Online]. Available: <https://www.techpolicy.press/right-capacity-and-will-in-content-moderation-a-case-for-user-empowerment/>
- [5] Steve Carey, "AI Villains: The Human Cost Behind Content Moderation and AI Development," LinkedIn, 2024. [Online]. Available: <https://www.linkedin.com/pulse/ai-villains-human-cost-behind-content-moderation-steve-wxqbc>
- [6] Richard J. Andreano et al., "Disparate Impact Executive Order and HUD to Reconsider Disparate Impact Rule," 2025. [Online]. Available: <https://www.consumerfinancemonitor.com/2025/05/02/disparate-impact-executive-order-and-hud-to-reconsider-disparate-impact-rule/>
- [7] Maxicus, "Social Media Content Moderation: The Complete Guide,". [Online]. Available: <https://maxicus.com/social-media-content-moderation/>
- [8] Gokan Ozcifci, "How to create a scalable Power Platform governance model for growing adoption," Syskit, 2025. [Online]. Available: <https://www.syskit.com/blog/scalable-power-platform-governance-guide/>
- [9] Iberdrola, "Digital rights, essential in the Internet age," Iberdrola Innovation. [Online]. Available: <https://www.iberdrola.com/innovation/what-are-digital-rights>
- [10] Emmanouil Papagiannidis et al., "Responsible artificial intelligence governance: A review and research framework," The Journal of Strategic Information Systems, Volume 34, Issue 2, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0963868724000672>
- [11] Deepesh PC, "Why governance in Power Platform is imperative for organizations," Coforge. [Online]. Available: <https://www.coforge.com/what-we-know/blog/why-governance-in-power-platform-is-imperative-for-organizations>