Volume **8** (3) 2026 https://doi.org/10.5204/lthj.4614

Law as a Game: Shaping Human Behavior in Virtual Worlds through Gamified Regulatory Logic

David Torabi
Kyushu University, Japan

Abstract

This article introduces gamified regulation as a legitimate mode of governance for immersive virtual reality (VR) worlds commonly described as the *Metaverse*. It begins with the observation that such environments are ontologically distinct and architected, enabling forms of regulation that are not merely imposed externally but embedded within the conditions of participation themselves. Building on this premise and drawing on examples from online games and existing virtual platforms, the article explores how game and related design mechanisms, such as reputation systems, play-based sanctions, participatory adjudication, AI-assisted enforcement, and blockchain-based rule execution, can be leveraged to shape behavior and (self-)sustain normative order from within the virtual environment. Rather than offering a fixed framework or prescriptive model, the analysis develops a conceptual lens for understanding and critically assessing these regulatory possibilities. While gamified regulation promises more proactive and experientially coherent governance that resonates with the game-like nature of the Metaverse, it also raises significant ethical, legal, and social concerns, including risks of trivializing harm, surveillance, vigilantism, algorithmic bias, social stratification, and the normalization of behavior-based control. The article therefore positions gamified regulation as a concept that expands the jurisprudential imagination, while requiring sustained scrutiny to ensure that fundamental human values are preserved in immersive digital societies. This scrutiny should start from the constitutional question of who authorizes the architects of such worlds, and the article proposes the republican ideal of freedom as non-domination as a guiding criterion.

***Keywords**:* Metaverse governance; gamification of law; legal pluralism; speculative jurisprudence; socio-legal theory; law reform research.

1. Introduction

Law has long traveled with humankind. Across legal traditions and political orders, it has often portrayed itself, somewhat paradoxically, as both part of and outside the stream of ordinary life, a structure of rules that watches, evaluates, condemns, and, when necessary, punishes. But in emerging immersive virtual worlds, often collectively associated with the phenomenon of the Metaverse, that familiar self-image arguably begins to lose clarity.

The Metaverse—as discussed in this article—consists of persistent, immersive, and socially inhabited virtual reality (VR) worlds that users experience as avatars through VR technology (head-mounted displays, special audio devices, haptic sensors, etc.). At its core, the Metaverse is a hybrid phenomenon that shares traits of both the physical world and conventional cyberspace. This hybridity is not merely descriptive but constitutive of its normative and experiential structure. On the one hand, the Metaverse inherits from the physical world a sense of persistence, spatiality, and embodied presence: users move through environments and experience interactions that are felt not as abstract outputs but as situated, often affectively charged events. On the other hand, it retains the malleability and infrastructural dependency characteristic of conventional cyberspace,

Except where otherwise noted, content in this journal is licensed under a Creative Commons Attribution 4.0 International Licence. As an open access journal, articles are free to use with proper attribution. ISSN: 2652-4074 (Online)

© The Author/s 2026

where everything that does or does not exist ultimately results from code and design choices. The Metaverse is thus phenomenologically and ontologically distinct as it dissolves, to an extent, the clear separation between lived physical reality and cyberspace as we know it, resulting in something we might call 'quasi-corporeality.' Consequently, (mis)conduct within such environments cannot be clearly classified as either 'real' or 'virtual,' nor should the normative frameworks governing it rely on that simplistic dichotomy. Instead, behavior unfolds in a space where physical intuitions about harm and responsibility intersect with digitally mediated architectures, producing a layered ontology in which actions are simultaneously embodied (experienced) and encoded (engineered). These attributes make the Metaverse a particularly interesting phenomenon for law reform research and theoretically grounded speculation, inviting us to reconsider not merely the reach of law, but its very form. The term 'Metaverse' itself requires some clarification. When Facebook named itself Meta in 2021, it posited the Metaverse as the next iteration of the internet, and much of the technology industry followed suit. Little of that proposition has materialized. In early 2026, Meta pivoted much of its focus from the Metaverse to a transformation around artificial intelligence (AI), while limiting support for its flagship social VR platform, *Horizon Worlds*.[1] Explaining the decision, Meta's chief technology officer described the Metaverse as a 'misunderstood concept' that was never meant to be only VR.[2] In concrete terms, the Metaverse's definition remains unsettled. Ambitious accounts such as Ball's require, among other things, persistence, live and synchronous experience without a cap on concurrent users, a fully functioning economy, and interoperability of data and assets across worlds, a combination that no platform existing in 2026 offers.[3] In ordinary use, the term stretches to cover very different things, from social VR platforms to online multiplayer games and blockchain-based applications.

For regulatory imagination, this lack of consensus carries two distinct notions. It invites speculation about an object that does not (yet) exist, and claims about 'regulating the Metaverse' say little unless it is clear which environments they mean. At the same time, because these environments have not settled on a dominant corporate design, we still have room to consider how they ought to be governed before particular architectures become entrenched. This article thus does not use the term 'Metaverse' in terms of provenance or prophecy, but as a shorthand for a cluster of features (notably, persistence, immersion, code-embodied presence, sociality, and rules coded in advance by designers) that already exists, unevenly and on a far smaller scale than once announced, most clearly in social VR platforms, and that may converge further in the future.

These features also set such environments apart from the online multiplayer games that supply many of this article's examples. Multiplayer games are often persistent and social, and their communities have long developed their own norms and sanctions. The boundary is not sharp, but the differences matter for regulation. In social VR, users are present through tracked body movements and simulated proximity, so users feel their likeness more directly than when a character is controlled across a screen.[4] Social VR spaces are, for the most part, not organized around game logic or mentality by default. People go there to socialize, so that misconduct is harder to dismiss as just playing a game, and sanctions affect a person's social life rather than a specific game session. Social VR may thus be game-like in its architecture, but not, or not only, a game in its purpose. In this article, games serve as a reservoir for regulatory sandbox experimentation, while the case for a new regulatory imagination rests on what happens when these experiments carry into spaces that are not games.

Within the horizon of this clarified premise, this article contributes primarily to speculative jurisprudence, legal pluralism, and socio-legal theory (most notably in the field of law reform research). It uses immersive virtual worlds as a conceptual and exploratory stress test for law itself and advances a socio-legal theory of the Metaverse as an ontologically distinct regulatory space in which gamification is a conceptually and normatively defensible response to immersive harm and governance. To situate the article more clearly, I borrow Arthurs's typology of legal research styles as a visual aid (as adapted by Chynoweth, see Figure 1 below). The dotted circle marks the secondary focus, the other the primary focus. This is, of course, not an exact placement, but an approximation of where the article best fits within legal scholarship. The examples used throughout illustrate the argument rather than test it empirically, and (empirical) research from the perspective of actual participants would be a valuable complement.

[1] Bobrowsky, "Meta Laying Off Thousands."
[2] Bell, "Meta Isn't Shutting Down."
[3] Ball, "The Metaverse."
[4] See Lin, "Digital Body"; Bellini, "Virtual Justice," 75.

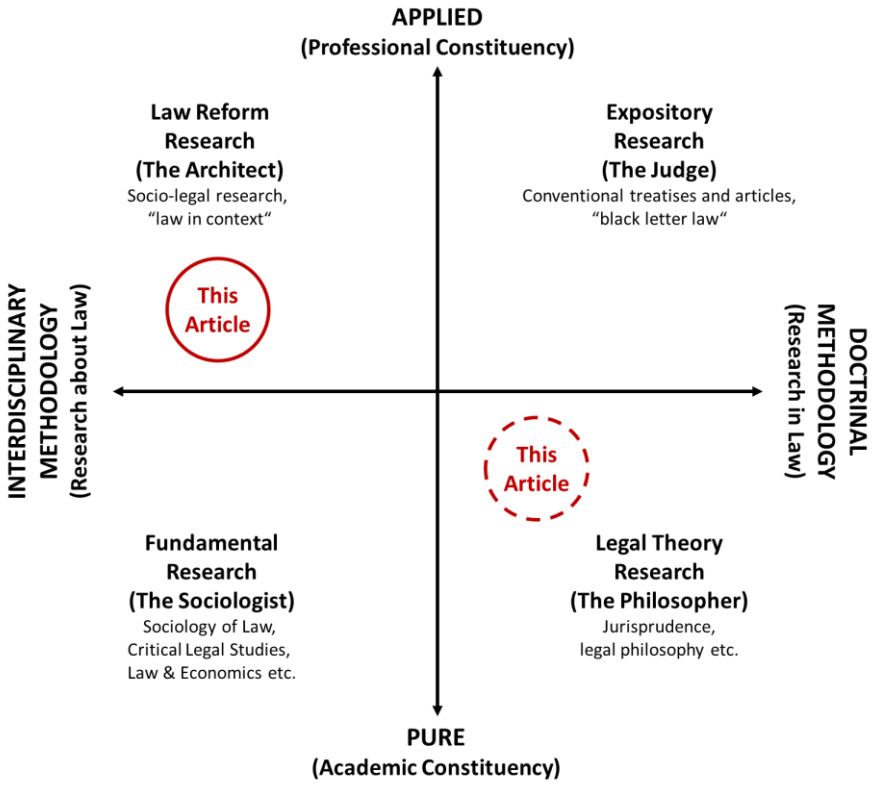


Figure 1. Categorization of Legal Research (after Arthurs 1983)[5]

The article also builds on, and hopes to contribute to, the growing scholarship on law and video games. Work in the law and humanities and cultural legal studies traditions has shown that games do not merely depict law but allow players to experience it. Newbery-Jones argues that video games, by calling for active participation rather than passive observation, are well suited to exploring justice as a phenomenological experience, and has proposed their use in experiential legal education because they let players experiment with ethical choices without real-world consequences.[6] Barnett and Sharp have sought to show how the narrative and mechanics of the video game *inFAMOUS* shape the player's understanding of law, morality, and power, arguing that the game privileges a natural law conception.[7] Mitchell, Pearson, and Peters extend these cultural legal perspectives by examining how players engage with legality-shaped game environments and how the rules coded into games and the rules of law inform one another.[8] This article argues that games are a serious site of legal meaning, but asks less about how law appears in games than about how game mechanics might themselves become a form of law in immersive virtual social societies, where conduct is not entirely free of real-world consequences.

The article does not claim that the Metaverse ought to be left to private ordering, or that criminal and public law lose their relevance in immersive environments. Its claim is more modest, but also more demanding. The Metaverse, precisely because it is architected and experientially distinct, makes possible a mode of governance qualitatively different from the legal techniques fashioned for conventional life, online or offline. In such environments, law need not appear only as a prohibition backed by retrospective sanction. It can be woven into the world's own social mechanics. Understood this way, gamified regulation may constitute a distinctive jurisprudential form.

[5] Slightly modified from Chynoweth, "Legal Research in the Built Environment," 671.
[6] Newbery-Jones, "Answering the Call of Duty"; Newbery-Jones, "Ethical Experiments with the D-Pad."
[7] Barnett, "The Moral Choice of inFAMOUS."
[8] Mitchell, Law, Video Games, Virtual Realities.

2. The Metaverse as an Architected Regulatory Space

Metaverse worlds are designed worlds, and design shapes conduct. Unlike social media and similar conventional forms of digital interaction, virtual worlds are not merely a form of communication, but a new variant of living. Yet they are not quite like ordinary physical life, because these worlds are built environments whose rules designers code in advance. Their social architecture, the rules and mechanics that determine what users can do and become, therefore deeply shapes interaction patterns. This is, in essence, Lawrence Lessig's well-known point that in digital environments 'code is law': alongside law, social norms, and the market, architecture regulates conduct, often more effectively, because its constraints operate in advance and need no enforcement after the fact.[9] Lessig's proposition is simultaneously a warning that regulation through code can escape the transparency and accountability expected of law, a concern that runs through the analysis below. Cultural legal scholarship on video games has added that players do not simply submit to the legality coded into games but engage with and interpret it, and that law itself can be read as a kind of game with its own roles and rules.[10] The architectures that arise within Metaverse worlds can support what Robert Cover called 'jurisgenesis,' the process through which the communities inhabiting them create law-like normative orders out of their own narratives and commitments, but at the same time they have the capacity to be jurispathic (i.e., capable of suppressing such orders and of destroying or undermining legal and moral principles).[11] To Cover, the community, not the institutional designer, is the progenitor of legal meaning, a point I will take up again in Section 4. Understanding this normative tension matters because, simply put, the architectures of the Metaverse can cultivate good behavior or enable bad behavior, depending on how we, as a society, comprehend and build them.

Consider the online game *EverQuest*. Designers introduced a mechanic according to which players who died lost their loot and had to return to their corpse to reclaim it. What looked at first like a punitive feature had a quite unexpected social effect: it encouraged altruism, because players often had to rely on others to recover their belongings.[12] A mechanic meant to increase difficulty also fostered cooperation and mutual aid. The game *Albion Online* offers a different illustration. In its full-loot 'Black zones,' players lose all their items if another player defeats them.[13] That design decisively shapes conduct, encouraging both predatory behavior and loss-aversion strategies. Such examples show that (game) design choices do not merely accompany social norms, but can help produce them.

Virtual worlds make available protective and sanctioning mechanisms that would be impossible, or at least far more limited, in the physical world. Because the Metaverse is not bound by the constraints of material reality, it can respond to misconduct in more imaginative ways. This became evident in *QuiVr*, a virtual reality game in which users reported sexual harassment. The developers responded by introducing a 'power gesture' superpower that let players perform certain hand motions to instantly create a personal force-field bubble around themselves.[14] Avatars entering that bubble were muted and rendered invisible, effectively removing them from the user's experiential space. The feature gave users immediate power to enforce boundaries in a way that is difficult to replicate offline.

More generally, designers can implement safe zones, collision rules, visibility filters, altered physics, and related tools to prevent or respond to virtual offenses. Meta's 'personal boundary' feature provides one such example, introducing a mechanism that prevents one avatar's outstretched hand from grabbing another without mutual consent.[15] The broader jurisprudential significance here is that justice in the Metaverse need not depend solely or primarily on external authorities. It can be enacted, at least in part, by the world itself.

3. From External Law to Internalized Governance

In 2008, Orin Kerr argued that criminal law, with all its bluntness, should be categorically excluded from virtual (then predominantly 3D) worlds regardless of how advanced they become, on the ground that it rarely offers a sufficiently nuanced response to social harms.[16] While this conclusion may be overstated, the concern underlying it warrants serious consideration. Criminal law is indeed a blunt instrument, and we should not transpose it into unfamiliar environments without asking whether it fits the medium's logic and whether the medium enables more proportionate responses.

[9] Lessig, Code: Version 2.0, especially chapters 1 and 7.
[10] Mitchell, Law, Video Games, Virtual Realities, see especially the contributions in Part I ('Code as Law, Law as Play').
[11] See Shapiro, "The Disappearance of Cyberspace," 720; Cover, "Nomos and Narrative."
[12] Sparrow, "How to Govern the Metaverse."
[13] See Albion Online Wiki, "Types of Regions."
[14] McAloon, "QuiVR Devs Created New Features."
[15] Sharma, "Introducing a Personal Boundary Horizon Worlds."
[16] Kerr, "Criminal Law in Virtual Worlds," 417, 425–28.

In a game-like environment such as the Metaverse, gamified sanctioning systems present one such response. They work *with* the medium's logic rather than *against* it. The distinctive features of virtual worlds support the view that Metaverse communities, rather than replicating offline criminal law in digital form, might draw on game-design mechanics to develop justice systems embedded within the virtual world itself. This is particularly relevant because players in online multiplayer games have often been observed to show apathy toward, or even resist, perceived external interventions in their virtual affairs.[17] For that reason, the Metaverse may, in certain respects, be governed more effectively through the gamification of regulation, leveraging the multifaceted, game-like architecture of virtual worlds to incentivize desirable conduct and sanction misconduct in ways better attuned to their context.

This approach's appeal lies not merely in its novelty. It lies in its congruence with the structure of immersive environments, and in its capacity to remedy the shortcomings of simply transplanting offline legal forms into digital spaces. Because gamification uses mechanics that intrinsically motivate players, a gamified regulatory model also offers a route away from reactive enforcement and toward genuinely proactive governance. Under such conditions, justice becomes part of the world's ordinary rhythm. More important still, it not only deters and corrects deviant behavior but also guides virtual societies toward shared norms through internal motivation rather than external fear.

However, it is also crucial to understand and respect that certain behaviors are inherently incompatible with gamification because of their severity. Thus, they should be reserved for conventional criminal law, if necessary, by extension or adaptation. The illustration below, the typology of which is based on the 2024 INTERPOL White Paper,[18] shows broadly to what extent gamification would be leveraged in the proposed regulatory concept (Figure 2). The illustration should be understood as a prototypical categorization and scope, not as a final placement.

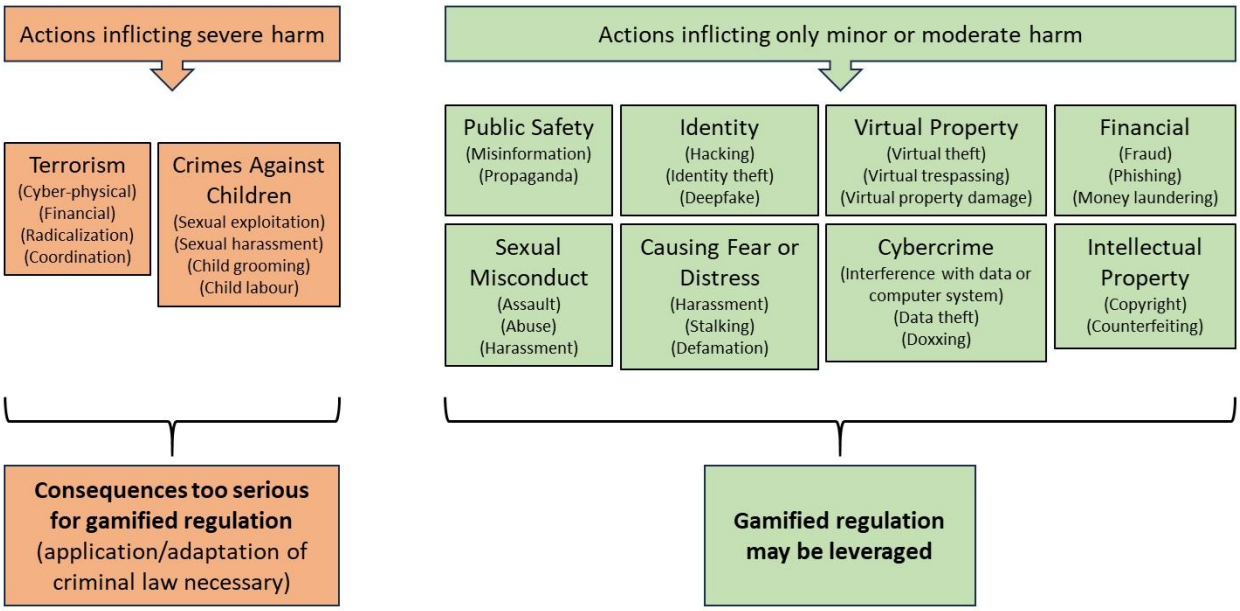


Figure 2. Categorization of Metaverse Misconduct and Scope of Gamification

[17] See Sparrow, "How to Govern the Metaverse."
[18] INTERPOL, Metaverse, 12–15.

4. Law as a Communal Resource in Virtual Worlds

At the center of Metaverse socio-legal theory lies a reimagining of law. Law should not be seen merely as an instrument of restraint or command, but as a communal resource that enables human agency and sustains collective life. This is not a paternalistic or technocratic ideal, but a view of law as an enabling structure, one that supports both individual and shared agency by clarifying expectations and embedding responsibility in the Metaverse's social fabric. In the right setting, agency can empower individuals and enable them to engage responsibly with their liberty.

Crucially, the conception of agency, and consequently liberty, at work here is to be understood not as modalities of egoist exclusivity. Drawing on a Rawlsian and Hohfeldian sensibility, agency here is relational. It arises through a network of rights, duties, privileges, and obligations that organize coexistence and cooperation.[19] If digital societies are to become (semi-)self-sustaining, ethical conduct must come to resemble something more like an internal disposition than an externally coerced performance. Law contributes to that outcome not chiefly by threatening punishment, but by shaping the architecture within which people act. Oversight and enforcement remain necessary, but should function as auxiliary means rather than first principles.

What matters here, fundamentally, is who should be responsible for shaping that process, which is where Cover's account of law, already mentioned in Section 2, is rather instructive. For Cover, law in the full sense, a *nomos*, does not arise from rules alone, but from communities that locate their rules in shared narratives and commit themselves to living by their own interpretation of them.[20] In a designed world, this relationship is easily inverted, because the platform designer writes the rules in advance and enforces them through code, and thus occupies the place Cover reserves for the community. Taken to the extreme, this is, at its core, jurispathy, because designers declare one normative order to be law and suppress the others.[21] Hobson's case study of *VRChat* finds that social VR spaces in which users moderate themselves and co-produce their own rules can form more robust, self-sustaining communities than spaces governed by platform-imposed rules, suggesting this is not merely theoretical but practical.[22] If law in the Metaverse is to be a genuine communal resource, the architecture must leave room for communities to generate, interpret, and if necessary revise norms on their own, rather than receiving them as presupposed features of the world they inhabit.

This naturally raises the problem of constitutional justification, which ties to the so-called 'God problem' (i.e., a state where designers have unchecked power within a given governance architecture) and underlies any process of designing law in the Metaverse.[23] At present, constitutional legitimacy is largely given in terms of service, which, as Suzor has noted for digital platforms more generally, gives operators broad discretion over users while leaving them with few enforceable rights. His response is to hold private platform governance to the values of the rule of law.[24] Teubner's work on societal constitutionalism offers a broader idea for the same problem, describing how private regimes beyond the state can develop their own constitutional fragments that both constitute and limit their governing powers, typically through self-limitation prompted by outside pressure.[25] Applied to virtual worlds, these ideas suggest the necessity for at least four elements that should guide the process of designing law in the Metaverse: (i) a clear distinction between the power to make the basic rules of a world (constituent power) and the powers exercised under those rules (constituted power); (ii) rights that participants hold against the platform, rather than privileges that the platform grants and can revoke; (iii) a separation of functions where one architecture would otherwise legislate, monitor, adjudicate, and execute at once; and (iv) procedures for amending the rules that are open to participants as members of the community, and not only as a reward for good conduct.

Law in the Metaverse can then become more than a list of prohibitions. It can become a living framework, collectively sustained, through which digital communities are bound by mutual recognition and normative clarity. In a Rawlsian spirit, such an order aspires to fairness by reflecting principles that participants could reasonably accept even under conditions of uncertainty. Participants must know where their freedoms end, where others' rights begin, and how responsibility is allocated within shared digital environments. Understood in this way, law becomes the largely invisible architecture of collective virtual life, enabling digital societies to flourish without domination. Domination here should be understood in the demanding sense that the republican tradition gives the term. For Pettit, a person is dominated when another has the capacity to interfere in their choices

[19] Rawls, A Theory of Justice, 21–29; Hohfeld, "Fundamental Legal Conceptions."
[20] See Cover, "Nomos and Narrative."
[21] See Cover, "Nomos and Narrative," 4–11, 53.
[22] Hobson, "Emergent Governance from Polycentric Order."
[23] See Grimmelmann, "Virtual Worlds as Comparative Law," 175–83.
[24] Suzor, Lawless.
[25] Teubner, Constitutional Fragments.

on an arbitrary basis, whether or not that capacity is ever exercised, so that even the slave of a kind master remains unfree.[26] Interference is not arbitrary when it follows the interests and views of those affected, which in practice requires that they can contest it and have a say in the rules governing it. This standard matters for the mechanisms of gamified law discussed in the following section, and thus they should be read with this theoretical backdrop in mind. Even a constructive system of game mechanics would remain a form of domination if the architect who runs it is accountable to no one, even if no gamified sanctions were ever actually imposed. The question to ask of each mechanism is not only whether it is effective or well-intentioned, but whether those it governs can control and contest the power behind it.

5. Conceptual Design of Gamified Law in Virtual Worlds

What follows develops a speculative, conceptually grounded jurisprudence for platform-based virtual societies, outlining how gamified sanctions and incentives might (self-)sustain order while fostering prosocial behavior in immersive environments. It does not seek to offer a ready-made regulatory blueprint. Rather, it encourages legal scholars and policymakers to approach emerging regulatory challenges, such as those posed by the Metaverse, with greater imaginative openness and anticipatory care. The concept is neither an endorsement nor a critique in itself, but aims to provide a sober, high-level assessment of both the potential and the risks inherent in gamified forms of governance.

5.1. Reputation as Gamified Governance in the Metaverse

In a gamified Metaverse justice framework, reputation emerges not merely as an auxiliary signal of trust, but as a structuring principle of social order, redistributing functions traditionally reserved for formal legal institutions into everyday interactions. Each user might carry a reputation score that operates as a persistent marker of past conduct, akin to a hybrid of criminal record and credit rating, yet more immediate in its visibility and more pervasive in its consequences. Good conduct builds reputational capital, while misconduct diminishes it, shaping both access to opportunities and others' perceptions of one's character. The normative implications here require closer scrutiny, particularly when reputation ceases to be merely descriptive and becomes constitutive of one's standing.

Rather than formal adjudication, negative actions such as harassment or fraud could be translated into quantifiable demerits, logged, processed, and made visible across platforms. Proposals for such systems already exist in discussions of Metaverse governance, including portable reputational identities that transcend individual virtual worlds.[27] The underlying assumption is that visibility produces deterrence, as knowing one's actions will be publicly recorded and socially legible may discourage misconduct through anticipated shame or reputational loss.[28] However, this logic of 'regulation by shaming' presupposes relatively stable, shared normative standards and a proportional relationship between conduct and reputational consequence. In heterogeneous, globalized Metaverse environments, these conditions are far from guaranteed. What counts as misconduct, how it is interpreted, and how severely it is penalized may vary across communities, raising the risk that reputation systems entrench not only accountability but also bias, exclusion, or cultural misunderstanding.

If negative conduct can be encoded as reputational loss, positive conduct can likewise be translated into visible markers of esteem. Trust badges, displayed on avatars or profiles, would signal virtues such as cooperation, reliability, or conflict mediation. In this sense, reputation becomes explicitly gamified: badges resemble achievements in conventional games, transforming ethical behavior into a system of rewards and symbolic capital. The promise is distributed governance. By incentivizing desirable conduct and publicly recognizing it, the system encourages users to internalize community norms, reducing reliance on external, centralized enforcement. Yet this transformation also raises deeper questions about moral motivation. When virtue is externalized into badges and scores, there is a risk that ethical action becomes instrumentalized, pursued not for its intrinsic value but for its reputational payoff.

Reputation may further structure access to the Metaverse itself, operating through tiered systems that condition participation on accumulated trust. High-reputation users might gain entry to advanced economic opportunities or governance roles, while those with lower reputations face exclusion or restriction. Existing platforms already gesture toward such architectures. The social VR platform *VRChat*, for example, uses a trust rank system that differentiates users from 'visitor' to 'trusted user,' granting privileges such as the ability to upload custom content only to those who have demonstrated sufficient engagement and reliability.[29] In this configuration, reputation functions simultaneously as a deterrent and reward mechanism, shaping not only behavior but also the distribution of power and opportunity within virtual worlds. The analogy to offline systems of

[26] See Pettit, Republicanism, 23, 32, 52.
[27] Bermejo Fernandez, "Life, the Metaverse and Everything," 272; Qin, "Identity, Crimes, and Law Enforcement."
[28] Johnson, "Regulation by Shaming," 1866; Shore, "To Share or Not to Share."
[29] See VRChat Inc., "VRChat Safety and Trust System."

citizenship or creditworthiness is apparent and instructive, but also cautionary. Just as credit scores can entrench socioeconomic inequality, reputational tiers in the Metaverse might solidify hierarchies that are difficult to contest or escape.

5.2. Play-Based Sanctions and Their Limits

A shift toward play-based deterrence in the Metaverse invites us to reconsider punishment not as an external imposition, but as an internally experienced modification of one's capacities within a shared environment. Instead of fines or incarceration, a gamified system would embed penalties directly into the conditions of participation. Temporary debuffs, mandatory quests, or restrictions on visibility and interaction would signal wrongdoing and also reshape how the offender can interact and be perceived within the world itself. This reconfiguration aligns with the broader premise that virtual environments collapse the distance between rule and experience, but caution is needed to avoid trivializing harm by translating it into the idiom of play.

The concept of 'debuffs,' drawn from game design, illustrates both the promise and the ambiguity of this approach. A debuff weakens a character's attributes, and in a Metaverse context, this logic could be extended to users who violate community norms. Rather than being excluded outright, an offender might experience diminished movement, reduced economic capability, impaired social interaction, and similar constraints that are felt continuously and situationally. For instance, a user found responsible for harassment could face a temporary reduction in communicative reach or expressive capacity, limiting their ability to repeat the harmful conduct. Existing gaming practices offer suggestive precedents. In the game *Grand Theft Auto Online*, persistent misconduct such as 'griefing' results in a 'Bad Sport' designation, marked visually by a 'dunce cap' and accompanied by confinement to restricted lobbies for a period of time.[30] The sanction is playful in its aesthetic, yet consequential in its effects, combining inconvenience and reputational stigma. However, the very playfulness that renders such penalties palatable may also undermine their normative clarity. If punishment is experienced as a game mechanic, users may interpret it as part of the system to navigate or even exploit, rather than as a moral response to wrongdoing. The deterrent effect thus depends not only on the severity of the debuff, but on whether users continue to see it as expressive censure rather than merely an obstacle.

The idea of 'mandatory quests' extends this logic by transforming punishment into structured participation. Instead of paying a fine or enduring passive restriction, offenders could be required to complete in-world tasks that benefit the community as a condition of restoring their standing. These tasks might involve assisting new users or generating content of shared value (like building or improving virtual environments), thereby linking sanction to restitution. The underlying intuition resonates with restorative justice practices, which emphasize repair and reintegration over exclusion. Notably, similar proposals have already been advanced in discussions of virtual governance, where alternatives to automated bans include community service or mentorship.[31] Even in existing gaming contexts, elements of this approach appear, such as when the president of Daybreak Game Company offered banned players the opportunity to return if they publicly acknowledged wrongdoing in the battle royale game *H1Z1* (now *Z1 Battle Royale*).[32] The Metaverse amplifies these possibilities by rendering such practices more immediate and embodied through its immersive technology. Yet the normative dissonance persists. When rehabilitation is gamified, repair can become instrumentalized performance, undertaken to regain status rather than acknowledge harm. The system may produce outward compliance while leaving underlying attitudes unchanged, raising doubts about whether it achieves genuine moral transformation or merely simulates it. This difficulty is, of course, not peculiar to gamified sanctions. It attends restorative practices in general, where apology and reparation may likewise be offered strategically to secure a better outcome.[33] Gamification may add that the reward for repair is built into the system and openly displayed, which can make an instrumental attitude more tempting.

A more subtle form of deterrence can be described as 'avatar exile,' where the offender's presence is not eliminated but attenuated. Instead of a total ban, the user remains within the environment but is rendered visually diminished, perhaps as a gray or shade-like figure, and functionally constrained in their interactions. This form of 'social dimming' makes exclusion perceptible without fully severing participation. The offender loses recognition and influence, encountering the community from a marginal position. Existing platforms already approximate this model. *VRChat*'s trust system assigns a 'Nuisance' rank to problematic users, marking them visibly and limiting how others perceive their avatars, curtailing their ability to affect shared spaces.[34] In more extreme cases, the analogy to the 'ghost state' in games such as *World of Warcraft* becomes apparent, where a user may be present but unable to meaningfully interact. This approach has a certain normative appeal: it avoids the finality of exclusion while making the consequences of misconduct experientially conspicuous. At the same time, it raises concerns

[30] Black, "Dunce Cap in GTA Online."
[31] See Sparrow, "How to Govern the Metaverse."
[32] Klepek, "H1Z1 Cheaters Publicly Apologizing."
[33] See, in general, Menkel-Meadow, "Restorative Justice"; Braithwaite, "Restorative Justice."
[34] See VRChat Inc., "VRChat Safety and Trust System."

about human dignity and proportionality, as being rendered invisible or socially irrelevant is a distinct form of symbolic degradation that may be deeply alienating.

5.3. The Politics of Incentivized Virtue

A mature conception of justice within the Metaverse cannot be confined to the reactive logic of sanction. Durable social spaces must encourage desirable conduct in advance, not only enforce it after the fact. Positive reinforcement thus becomes a central regulatory modality, stabilizing norms by rewarding cooperation, responsibility, and prosocial engagement before more serious harms arise. The aim is to reward virtue without reducing it to strategic optimization within an incentive system.

One manifestation of this approach could be vigilante titles and in-world bounties, where users are rewarded for helping detect and mitigate harmful conduct. Just as games incentivize quest completion, Metaverse platforms might grant advantages (e.g., tokens or points) to users who successfully report infractions. This could give rise to forms of crowdsourced moderation in which community members assume quasi-enforcement roles, accumulating titles such as 'Community Overseer,' '[Metaverse Name] Guardian,' and the like in recognition of their contributions. These titles may be stratified, reflecting varying degrees of engagement and reliability and creating a hierarchy of informal authority within the user base. Existing gaming practices offer preliminary analogs, including systems that reward sustained good behavior with in-game benefits or monetary bonuses.[35] In the Metaverse, such mechanisms could evolve into micro-reward structures that embed enforcement within the community itself. This model is certainly not without risk, because it can incentivize users to monitor and report one another, encouraging over-reporting, strategic accusations, or undesirable digital vigilantism that undermines trust rather than reinforcing it. The line between communal responsibility and opportunistic enforcement is thus rather thin, particularly when reputational or material rewards are at stake.

A more structurally ambitious extension of positive reinforcement is the use of 'governance tokens,' particularly within platforms influenced by decentralized or Web3 architectures. Here, exemplary conduct is not recognized symbolically but translated into formal participatory power. Users who demonstrate sustained responsibility and adherence to community norms might be awarded tokens that grant voting rights over platform decisions or access to privileged features.[36] In this way, ethical participation becomes a pathway to political influence, echoing the governance structures of decentralized autonomous organizations (DAOs), where token holders collectively shape the system's rules.[37] The appeal of this model lies in its apparent alignment between virtue and authority: those who contribute most constructively are entrusted with greater responsibility for the community's future. However, this alignment is neither inevitable nor unproblematic. Token accumulation may privilege those with more time or resources, reproducing inequalities under the pretense of merit. These problems also exist in the real (physical) world, but they appear in the Metaverse in a new context and require scrutiny. Moreover, translating ethical behavior into governance power risks conflating personal moral values with political legitimacy, which legal and political theory should treat with caution.[38] The question is not only who governs, but on what basis their authority is justified, and whether a system of incentivized virtue can sustain the pluralism and contestation that characterize healthy democratic orders. Tokens awarded for good conduct are, moreover, a behavioral incentive. They cannot take the place of a political standing that participants hold as members of the community, prior to and independent of their conduct (see Section 4 above).

Complementing these mechanisms, the concept of 'buffs' for good actors exemplifies how positive reinforcement might be embedded directly into the experiential structure of the Metaverse. As the inverse of debuffs, buffs enhance a user's capacities or visibility, rewarding those who consistently contribute constructively to the community. Such enhancements could be conferred automatically, for instance through AI-driven moderation systems that identify patterns of prosocial behavior (e.g., users who de-escalate conflicts or provide assistance) and translate these patterns into increased trust levels or functional advantages. Existing systems in online games, including peer endorsement mechanisms that reward sportsmanship and cooperation, suggest that such approaches can reduce toxic behavior and promote more harmonious interaction.[39] In a Metaverse context, these dynamics could intensify through real-time recognition, such as immediately attributing reputational gains or visible distinctions when a user successfully defuses a conflict. Yet the apparent neutrality of such systems should not be taken for granted. AI-mediated recognition is inevitably shaped by its training data and the values embedded in its design. As a result, bias or discrimination may shape who is deemed deserving of reward, potentially reinforcing existing inequalities or marginalizing certain forms of expression and interaction.

[35] Black, "Dunce Cap in GTA Online."
[36] See Kraken Learn Team, "What is a Governance Token."
[37] Werbach, Stablecoin Toolkit, 21; Han, "A Review of DAO Governance."
[38] See, in general, Nagel, "Moral Conflict."
[39] See Sparrow, "How to Govern the Metaverse."

5.4. Participatory Justice in Virtual Societies

A gamified justice system within the Metaverse need not be conceived as a purely hierarchical apparatus imposed from above. Rather, it can be structured as a participatory order in which the community itself assumes an active role in adjudication. By embedding adjudicative functions within the user base, such systems could promise scalability and responsiveness, particularly in environments characterized by vast and rapidly evolving interactions. Yet they also raise difficult questions about when collective judgment can approximate justice rather than devolve into a majoritarian impulse.

The most direct expression of this participatory model is the notion of player juries and virtual trials, in which users review allegations of misconduct and determine appropriate outcomes. In such a system, reported violations could be presented, either anonymously or with limited identifying information, to panels of users selected for demonstrated reliability or high reputation. These users would deliberate, assess available evidence such as chat logs or behavioral data, and decide both whether an infraction occurred and what form of response should follow. Such a model would distribute moderation burdens and enhance perceived legitimacy. Peer-rendered decisions, after sufficient deliberation, may feel more contextually informed and socially grounded than those imposed by distant platform authorities. A past example, albeit a failed one, is Riot Games' 'The Tribunal' for the online game *League of Legends*, which allowed qualified players to evaluate reported misconduct and vote on sanctions based on in-game evidence.[40] Although the system was ultimately discontinued, it showed that decentralized adjudication can function, at least temporarily, as a mechanism for collective norm enforcement. For such a system to be feasible, it would need to address whether participants have the necessary impartiality and whether procedural safeguards, such as the presumption of innocence or the right to contest evidence, can be meaningfully preserved in such settings.

Closely connected to this participatory dimension would be the development of 'tiered offense systems,' which introduce gradations of wrongdoing and corresponding proportionality in response. Just as offline legal systems distinguish between minor infractions and serious crimes, a Metaverse justice framework could classify violations along a spectrum, ranging from low-level disruptions to severe forms of harm (within the limitations illustrated by Figure 2 above). Minor or first-time offenses might result in warnings or limited reputational penalties, while repeated or more serious conduct could trigger escalating responses, including feature restrictions, intensified debuffs, mandatory restorative tasks, temporary avatar exile, or eventual expulsion from the platform. By calibrating responses to the severity and context of behavior, such a system avoids the rigidity of one-size-fits-all sanctions and helps users more clearly understand the consequences of their actions. Moreover, when these gradations are rendered visible within the environment (through changes in avatar appearance, access, or capability), they become experientially legible, reinforcing the connection between conduct and consequence.
Existing efforts to categorize harmful behavior in online environments could provide a foundation for such systems. The *Disruption and Harms in Online Gaming Framework* developed by the Fair Play Alliance and the Anti-Defamation League's Center for Technology and Society offers a taxonomy of problematic in-game conduct that could be adapted to Metaverse contexts.[41] By distinguishing between different forms and intensities of harm, such frameworks enable more precise and proportionate responses. Yet here too, the normative stakes are significant. The classification of conduct is never truly neutral as it reflects underlying judgments about what counts as harm and how competing values, such as freedom of expression and protection from abuse, are balanced.

5.5. Automated Enforcement and the Problem of Algorithmic Justice

Automated systems are likely to become integral to Metaverse governance, not simply as auxiliary tools but as constitutive elements of how norms are monitored, interpreted, evaluated, and enforced. The density of immersive social worlds makes purely human moderation increasingly untenable, pushing us toward AI that can observe behavior, identify patterns, and intervene where appropriate. This development reflects a general trajectory in digital governance, where automation brings efficiency and responsiveness.[42]

In immersive environments, AI-powered surveillance and moderation could operate continuously, analyzing avatar behavior and patterns to detect emerging forms of harm. While such systems already exist in more limited forms on the conventional internet, their significance is amplified in virtual worlds where interactions are not merely textual but embodied and spatially manifested. Existing research suggests that algorithms can identify indicators of harmful conduct, such as offensive language or violations of personal virtual space that may signal harassment or escalating aggression.[43] If predefined thresholds are reached, such systems might intervene before misconduct fully materializes. This anticipatory logic gestures toward a form of

[40] See League of Legends Wiki, "The Tribunal."
[41] Fair Play Alliance, Disruption and Harms.
[42] See OECD, Governing with Artificial Intelligence.
[43] Babaeianjelodar, "Interpretable and High-Performance Hate," 233; Wang, "HardenVR," 94; Bellini, "Virtual Justice," 75.

preemptive governance that seeks to prevent harm rather than merely respond to it. However, this very orientation makes it normatively contentious. The prospect of intervening before a violation is complete evokes anxieties about pre-crime, where individuals are sanctioned not for what they have done but for what they are predicted to do. Yet dismissing such systems outright would overlook their potential to mitigate harm in environments where the consequences of misconduct are immediate and morally significant. The challenge, therefore, is not whether such systems should exist, but under what conditions they should operate.

Accordingly, automated systems may structure penalty escalation through mechanisms that track behavior over time and adjust responses accordingly. By maintaining records of prior conduct, AI could assign progressively more severe consequences to repeat offenders, creating a predictable enforcement trajectory. A first infraction might trigger a warning or minor restriction, while subsequent or more severe violations could lead to more substantial consequences and require additional human oversight (see Section 5.4 above on 'tiered offense systems'). Elements of progressive enforcement have appeared in existing platforms, such as *League of Legends*, where minor misconduct could be addressed through automated penalties while more serious or repeated offenses were subject to human review.[44] The underlying rationale draws on established principles of deterrence: when consequences are predictable and clearly communicated, individuals may be less inclined to reoffend.[45]

However, the apparent rationality of such systems should not conceal their normative complexity. Automated escalation presumes that past behavior reliably predicts future conduct and that predefined thresholds can operationalize proportionality. In practice, these assumptions may obscure contextual nuances in individual cases, including the possibility of misunderstanding or situational variation. Furthermore, reliance on algorithmic tracking introduces risks of opacity (AI 'black box' issue) and limited contestability, because users may not fully understand how decisions are made or have meaningful avenues to challenge them. Hence, the system's legitimacy hinges not only on its effectiveness in reducing harm, but equally on its capacity to remain transparent and subject to human review.

5.6. Utilizing Blockchain Architecture

In Metaverse environments shaped by Web3 infrastructures, enforcing norms directly through code further transforms the relationship between law, technology, and conduct. Rules are embedded within the very architecture of interaction, operating automatically and, in principle, without discretion.

Smart contracts exemplify this mode of auto-enforcement. As blockchain-based programs that execute predefined conditions, they can structure a wide range of Metaverse interactions, from economic exchanges to access arrangements and organizational memberships.[46] If this mechanism were extended to governance, misconduct detected (potentially by AI systems, see Section 5.5 above) could trigger immediate and automatic consequences, such as the restriction of access or the modification of privileges. In this configuration, enforcement becomes both instantaneous and ostensibly impartial, as the code generally applies rules without regard to identity or status. This may be particularly attractive in cases of financial wrongdoing, where delays or discretionary gaps could enable further harm. These features, however, also make such systems somewhat normatively complex. Smart contracts simply execute but do not interpret. Their rigidity presupposes that relevant contingencies can be exhaustively anticipated and encoded in advance, an assumption that sits uneasily with the variability and context-dependence of human behavior. Moreover, the appearance of impartiality may obscure the fact that particular actors design these systems, embedding specific values or priorities into code that may be difficult to contest once deployed. Where AI-based detection and automatic execution are combined in this way, a single technical system monitors, judges, and enforces at once, without the separation of functions discussed at the end of Section 4.

Integrating behavioral conditions into digital assets further illustrates how enforcement might become inseparable from the objects of virtual life themselves. Virtual goods (items, non-fungible tokens, avatars, etc.) could be programmed to respond dynamically to the conduct of their owners. A valuable asset might become temporarily inaccessible if its holder's reputation falls below a certain threshold, or a membership token might be suspended during a sanction. In this way, economic incentives align directly with normative expectations, creating immediate, tangible consequences for misconduct. The analogy to offline practices such as asset forfeiture comes to mind but is incomplete because, unlike traditional sanctions, which are imposed externally, these constraints are built into the object from the outset, making compliance a condition of continued use rather than a response to adjudicated wrongdoing. While such arrangements may enhance deterrence, they also presuppose the existence of a reliable and widely accepted metric of behavior, essentially an 'on-chain morality' that translates complex social judgments into quantifiable thresholds. This presupposition is relatively problematic, because behavioral evaluation is rarely

[44] See League of Legends Wiki, "The Tribunal."
[45] Gómez-Quintero, "A Scoping Study of Crime Facilitated by the Metaverse"; Johnson, "Regulation by Shaming," 1866; Shore, "To Share or Not to Share."
[46] See IBM, "What Are Smart Contracts."

neutral, and embedding it in asset functionality risks hardening contested norms into technical constraints. Opt-in governance structures may mitigate some of these concerns by allowing communities to consent to such arrangements, but power or information asymmetries may still shape consent.

More broadly, blockchain-based governance models open the possibility of participatory lawmaking within the Metaverse, where users contribute to creating and revising the rules that govern their interactions. DAOs provide a template for such systems, enabling token holders to vote on policy changes and shape the platform's evolution. Transposed into a Metaverse context, this could allow users, particularly those who have demonstrated sustained engagement or responsibility, to participate in defining offense categories, adjusting sanction parameters, or revising terms of service. Existing platforms such as *Decentraland* already point to this possibility, allowing governance token (Voting Power) holders to vote on proposals affecting world rules and platform features.[47] This model appears to align closely with ideals of participatory governance, rendering law adaptive and responsive to the community it regulates, but it may also concentrate influence in the hands of those with greater resources, potentially reproducing inequalities under the guise of decentralization (see Section 5.3 above on 'governance tokens').

5.7. Plural Governance through Virtual Zoning

In Metaverse environments composed of multiple spaces and heterogeneous communities, governance is not necessarily uniform, and it need not be. A gamified system may instead differentiate zones with varying levels of regulation, letting users navigate environments that prioritize safety, freedom, or experimentation to different degrees. One region might be a tightly regulated 'safe zone,' another a 'Wild West' where constraints are relaxed, and users bear more of their own risk. Player-versus-player online games have long used differentiated regions, such as the Blue, Yellow, Red, and Black zones in *Albion Online*, each with distinct rules on conflict and consequences.[48] Similarly, platforms like *Second Life* have used rating systems that segment spaces into General, Moderate, and Adult categories, structuring participation around content and behavioral expectations.[49] Zone-based governance thus appears to reconcile competing priorities by offering users choice, in line with a pluralist view that a single, uniform standard may not suit diverse communities.

That said, the idea that users can meaningfully choose between safety and risk presupposes a level of informed consent that may be difficult to achieve. Users may not fully understand the implications of entering less regulated zones, particularly in immersive environments where harms (ranging from psychological or social to economic threats) can be experienced with heightened intensity. Moreover, depending on the coded architecture, zone boundaries may not be as impermeable as they appear. Conduct originating in one space may have repercussions in another, and actors may exploit regulatory asymmetries by operating in loosely governed zones while targeting users from more protected environments. In such cases, fragmented governance risks undermining the baseline of order it seeks to preserve. In addition, zone-based governance may stratify Metaverse worlds into zones with differing rules, safety, and dignity, raising questions about the limits of voluntariness and the conditions under which consent can legitimize exposure to harm.

Nevertheless, zone-based governance remains compelling for accommodating diversity within expansive virtual worlds. It reflects the understanding that regulation in such environments cannot be monolithic, and that spatial differentiation may pragmatically balance competing values.[50] The challenge lies in ensuring that this differentiation does not devolve into fragmentation that leads to inequality and injustice, and that a core set of protections is maintained across all zones.

5.8. Gamification of Restorative Mechanisms

A gamified model of justice must also allow for restoration, because wrongdoing, while warranting a response, should not exhaust the offender's moral identity.

One approach to rehabilitation is empathy training (or similar) simulations that leverage the immersive capacities of virtual environments to foster moral understanding. By placing offenders in the perspective of those they have harmed, such simulations make the moral significance of wrongdoing vivid and tangible. A user sanctioned for harassment, for instance, might undergo a scenario in which they experience harassment directed at themselves, or observe a dramatized account of the emotional impact of their actions. Research on virtual embodiment supports the plausibility of such interventions, suggesting that first-person experiences in virtual reality can enhance empathy and alter perception. Early studies by Slater and others, as

[47] See the Voting Power (VP) delegates on the Decentraland website, accessible at https://decentraland.org/governance.
[48] Albion Online Wiki, "Types of Regions."
[49] Linden, "Maturity Ratings."
[50] See also Ball, "The Metaverse."

well as subsequent research involving domestic violence offenders, indicate that embodying another perspective can increase sensitivity to fear and improve recognition of harm.[51] These findings suggest that immersive technologies may offer moral education tools that are more immediate and more powerfully effective than traditional forms of sanction. Yet their use also raises concerns. The imposition of such simulations as a condition of rehabilitation introduces questions about autonomy and consent, particularly when individuals are compelled to undergo emotionally and psychologically intense immersive experiences.

Importantly, a system that permanently marks users with infraction histories or degraded reputational status risks creating a class of digitally stigmatized individuals, whose capacity for reintegration is effectively foreclosed. A more defensible model would allow for the gradual restoration of standing through sustained good behavior and meaningful engagement in rehabilitative practices. Over time, infractions might be expunged and full participation restored, provided the individual consistently demonstrates commitment to community norms. This approach aligns with long-standing restorative justice principles, which emphasize reintegration over exclusion and recognize the moral importance of forgiveness.[52] At the same time, it must contend with the complexities of public naming and shaming. While visibility of wrongdoing can serve deterrent and expressive functions, it also risks disproportionate harm, particularly when stigma persists beyond the period of sanction.[53]

6. Conclusion

The Metaverse does not simply extend the reach of law into a new domain. Rather, it compels us to reconsider what law is and how it is experienced. In immersive virtual environments, governance can be embedded in the architecture of the world itself, operating through incentives and consequences encountered as part of lived experience. This opens the possibility of new regulatory logics that are not feasible in the physical world, and gamified governance is among the most compelling.

Yet this transformation should not be celebrated uncritically. The same (game) mechanisms that make governance more engaging and responsive can also intensify control in subtle and far-reaching ways. As the preceding concept analysis has shown, systems built on algorithmic reputation, behavioral data, and participatory enforcement pose several challenges and risks. Reputation metrics may evolve from tools of trust into instruments of exclusion, reinforcing social hierarchies and conditioning interaction through reductive indicators of worth.[54] Meanwhile, the consensus among scholars, as of writing this article, seems to be that algorithmic infrastructures sustaining such systems remain prone to bias, arbitrariness, disproportionate surveillance, and limited contestability.[55]

The cumulative effect of these developments is a profound conflict at the heart of gamified Metaverse justice. If created without scrutiny and left unchecked, Metaverse architecture could, over time, evolve into systems of conformity and pervasive control akin to social credit regimes.[56] It is therefore crucial, when creating these architectures, to distinguish between configurations that cultivate a sensible balance between genuine liberty and responsibility and those that reduce social life to a quantified regime of behavioral optimization.

The republican conception of non-domination offers a criterion for drawing that distinction. What matters is not how benign a mechanism appears, or how rarely its sanctions are used, but whether the power behind it is checked. A configuration is defensible to the extent that the rules on reputation, sanctions, and rewards are made and amended through procedures in which participants take part as members of the community; that rule-making, monitoring, adjudication, and enforcement are not merged in a single opaque system controlled by the platform; that those affected can contest decisions before a body other than the platform acting in its own cause, with meaningful human review; and that certain protections hold against the platform itself, as rights rather than revocable privileges. Where these conditions are absent, even a well-meant reputation system leaves users under the architect's arbitrary power, and under such conditions a drift toward social credit-like control is most likely.
If law is to function as a shared communal resource within immersive virtual worlds, it must remain oriented toward values that transcend pure game logic, lest pluralism and relational freedom be reduced to mere instruments of efficiency or

[51] Slater, "First Person Experience of Body Transfer"; de Borst, "First-Person Virtual Embodiment"; Seinfeld, "Offenders Become the Victim."
[52] See Menkel-Meadow, "Restorative Justice," 161; Braithwaite, "Restorative Justice," 323.
[53] Fritz, "Online Shaming," 686; see also Billingham, "Enforcing Social Norms," 997.
[54] Chen, "People Are Way Too Obsessed with Rank," 925.
[55] See, for example, Lin, "Digital Body"; Charles, "Artificial Intelligence for Data-Driven Decision-Making"; Gomez, "Algorithmic Arbitrariness."
[56] See Kühnreich, "Social Control 4.0?"; Adelmant, "Social Credit in China."

engagement. This requires a jurisprudence that is attentive not only to outcomes, but to the conditions under which those outcomes are produced.

The analysis also has implications for the study of law more generally, which I can only indicate here. Immersive virtual worlds show in concentrated form what is easy to overlook offline: that much regulatory work is done by architecture and incentives rather than by rules and sanctions, and that law is not only known but also experienced. Once regulation moves into design, rule-making, monitoring, adjudication, and enforcement can quickly merge into a single technical system. These observations reach well beyond virtual worlds, for instance, to the growing use of algorithmic systems in public administration, but addressing them properly is a task for future work.

If used with care, the gamification of law may enable viable new forms of governance in virtual worlds that could, in time, inform and improve existing regulatory models in other contexts. In that sense, the Metaverse might serve as a multifaceted testbed or sandbox for experimenting with novel approaches such as the gamified regulatory concept explored in this article. Without such care, however, it risks creating environments in which metrics and misguided morals govern individuals. The task ahead is therefore to remain attentive to the evolving Metaverse phenomenon and keep an open but critical mind regarding its potential forms of governance.

AI Disclosure Statement

No Artificial Intelligence Generated Content has been used for this article beyond basic AI-assisted copy editing (checks of grammar, spelling, and punctuation).

Bibliography

Adelmant, Victoria. "Social Credit in China: Looking beyond the 'Black Mirror' Nightmare." Center for Human Rights and Global Justice at NYU School of Law, April 2021. https://chrgj.org/2021-04-20-social-credit-in-china-looking-beyond-the-black-mirror-nightmare.

Albion Online Wiki. "Types of Regions." Accessed October 4, 2026. https://wiki.albiononline.com/wiki/Types_of_Regions.

Babaeianjelodar, Marzieh, Gurram Poorna Prudhvi, Stephen Lorenz, Keyu Chen, Sumona Mondal, Soumyabrata Dey and Navin Kumar. "Interpretable and High-Performance Hate and Offensive Speech Detection." In *International Conference on Human-Computer Interaction*, 233–44. Springer Nature Switzerland, 2022. https://doi.org/10.1007/978-3-031-21707-4_18.

Ball, Matthew. "The Metaverse: What It Is, Where to Find It, and Who Will Build It." Epyllion, January 13, 2020. Accessed October 4, 2026. https://www.matthewball.co/all/themetaverse.

Barnett, Michael and Cassandra Sharp. "The Moral Choice of inFAMOUS: Law and Morality in Video Games." *Griffith Law Review* 24, no 3 (2015): 482–99. https://doi.org/10.1080/10383441.2015.1126399.

Bell, Karissa. "Meta Isn't Shutting Down Its VR Metaverse After All." *Engadget*, March 19, 2026. https://www.engadget.com/ar-vr/meta-isnt-shutting-down-its-vr-metaverse-after-all-165520696.html.

Bellini, Olivia. "Virtual Justice: Criminalizing Avatar Sexual Assault in Metaverse Spaces." *Mitchell Hamline Law Review* 50, no 1 (2024): 75–110.

Bermejo Fernandez, Carlos and Pan Hui. "Life, the Metaverse and Everything: An Overview of Privacy, Ethics, and Governance." In *2022 IEEE 42nd International Conference on Distributed Computing Systems Workshops (ICDCSW)*, 272–77. IEEE, 2022. https://doi.org/10.1109/ICDCSW56584.2022.00058.

Billingham, Paul and Tom Parr. "Enforcing Social Norms: The Morality of Public Shaming." *European Journal of Philosophy* 28, no 4 (2020): 997–1016. https://doi.org/10.1111/ejop.12543.

Black, Chris. "Dunce Cap in GTA Online: All You Need to Know." *Sportskeeda*, October 11, 2021. https://www.sportskeeda.com/gta/dunce-cap-gta-online-all-need-know.

Bobrowsky, Meghan and Raffaele Huang. "Meta Begins Laying Off Thousands of Employees as It Transforms around AI." *Wall Street Journal*, May 20, 2026. https://www.wsj.com/tech/ai/meta-begins-laying-off-thousands-of-employees-as-it-transforms-around-ai-3c31a9eb.

Braithwaite, John. "Restorative Justice." In *The Handbook of Crime and Punishment*, edited by Michael Tonry, 323–44. Oxford University Press, 1998. https://doi.org/10.1093/oso/9780195110661.003.0013.

Charles, Vincent, Nripendra P. Rana and Lemuria Carter. "Artificial Intelligence for Data-Driven Decision-Making and Governance in Public Affairs." *Government Information Quarterly* 39, no 4 (2022): 101742. https://doi.org/10.1016/j.giq.2022.101742.

Chen, Qijia, Jie Cai and Giulio Jacucci. "People Are Way Too Obsessed with Rank: Trust System in Social Virtual Reality." *Computer Supported Cooperative Work* 33, no 4 (2024): 925–57. https://doi.org/10.1007/s10606-024-09498-7.

Chynoweth, Paul. "Legal Research in the Built Environment: A Methodological Framework." In *International Conference on Building Education and Research (BEAR): Building Resilience*, edited by Richard Haigh and Dilanthi Amaratunga, 670–80. Research Institute for the Built and Human Environment, University of Salford, 2008.

Cover, Robert M. "The Supreme Court, 1982 Term—Foreword: Nomos and Narrative." *Harvard Law Review* 97, no 1 (1983): 4–68. https://doi.org/10.2307/1340787.

De Borst, Aline W., Maria V. Sanchez-Vives, Mel Slater and Beatrice de Gelder. "First-Person Virtual Embodiment Modulates the Cortical Network that Encodes the Bodily Self and Its Surrounding Space during the Experience of Domestic Violence." *eNeuro* 7, no 3 (2020): ENEURO.0263-19.2019. https://doi.org/10.1523/ENEURO.0263-19.2019.

Fair Play Alliance and ADL Center for Technology and Society. *Disruption and Harms in Online Gaming Framework*. ADL Research Report, December 2020. https://thrivingingames.org/wp-content/uploads/2020/12/FPA-Framework.pdf.

Fritz, James. "Online Shaming and the Ethics of Public Disapproval." *Journal of Applied Philosophy* 38, no 4 (2021): 686–701. https://doi.org/10.1111/japp.12510.

Gomez, Juan Felipe, Caio Machado, Lucas Monteiro Paes and Flavio Calmon. "Algorithmic Arbitrariness in Content Moderation." In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 2234–53. ACM, 2024. https://doi.org/10.1145/3630106.3659036.

Gómez-Quintero, Juliana, Shane D. Johnson, Hervé Borrion and Samantha Lundrigan. "A Scoping Study of Crime Facilitated by the Metaverse." *Futures* 157 (2024): 103338. https://doi.org/10.1016/j.futures.2024.103338.

Grimmelmann, James. "Virtual Worlds as Comparative Law." *New York Law School Law Review* 49, no 1 (2005): 147–84.

Han, Jungsuk, Jongsub Lee and Tao Li. "A Review of DAO Governance: Recent Literature and Emerging Trends." *Journal of Corporate Finance* 91 (2025): 102734. https://doi.org/10.1016/j.jcorpfin.2025.102734.

Hobson, Anne. "Emergent Governance from Polycentric Order in Virtual Reality Social Spaces." In *Law, Video Games, Virtual Realities: Playing Law*, edited by Dale Mitchell, Ashley Pearson and Timothy D. Peters, 74–95. Routledge, 2024. https://doi.org/10.4324/9781003197805-5.

Hohfeld, Wesley Newcomb. “Fundamental Legal Conceptions as Applied in Judicial Reasoning.” *Yale Law Journal* 26, no 8 (1917): 710–70.
IBM. “What Are Smart Contracts on Blockchain.” IBM Think. Accessed October 4, 2026. https://www.ibm.com/think/topics/smart-contracts.
INTERPOL. *Metaverse: A Law Enforcement Perspective*. INTERPOL, January 2024. https://www.interpol.int/content/download/20828/file/Metaverse%20-%20a%20law%20enforcement%20perspective.pdf.
Johnson, Matthew S. “Regulation by Shaming: Deterrence Effects of Publicizing Violations of Workplace Safety and Health Laws.” *American Economic Review* 110, no 6 (2020): 1866–1904. https://www.aeaweb.org/articles?id=10.1257/aer.20180501.
Kerr, Orin S. “Criminal Law in Virtual Worlds.” *University of Chicago Legal Forum* (2008): 415–30.
Klepek, Patrick. “H1Z1 Cheaters Publicly Apologizing for Chance to Get Unbanned.” *Kotaku*, May 20, 2015. https://kotaku.com/h1z1-cheaters-publicly-apologizing-for-chance-to-get-un-1705865079.
Kraken Learn Team. “What Is a Governance Token?” Kraken, September 25, 2024. https://www.kraken.com/learn/what-is-a-governance-token.
Kühnreich, Katika. “Social Control 4.0? China’s Social Credit Systems.” *Eurozine*, August 10, 2018. https://www.eurozine.com/social-control-4-0-chinas-social-credit-systems.
League of Legends Wiki. “The Tribunal.” Accessed October 4, 2026. https://leagueoflegends.fandom.com/wiki/The_Tribunal.
Lessig, Lawrence. *Code: Version 2.0*. Basic Books, 2006.
Lin, Jinghuai and Marc Erich Latoschik. “Digital Body, Identity and Privacy in Social Virtual Reality: A Systematic Review.” *Frontiers in Virtual Reality* 3 (2022): 974652. https://doi.org/10.3389/frvir.2022.974652.
Linden, Jeremy. “Maturity Ratings.” Second Life Community Knowledge Base, February 24, 2011. https://community.secondlife.com/knowledgebase/english/maturity-ratings-r52.
McAloon, Alissa. “QuiVR Devs Created New Features to Combat VR Sexual Harassment.” *Game Developer*, October 26, 2016. https://www.gamedeveloper.com/design/-i-quivr-i-devs-created-new-features-to-combat-vr-sexual-harassment.
Menkel-Meadow, Carrie. “Restorative Justice: What Is It and Does It Work?” *Annual Review of Law and Social Science* 3, no 1 (2007): 161–87. https://doi.org/10.1146/annurev.lawsocsci.2.081805.110005.
Mitchell, Dale, Ashley Pearson and Timothy D. Peters, eds. *Law, Video Games, Virtual Realities: Playing Law*. Routledge, 2024. https://doi.org/10.4324/9781003197805.
Nagel, Thomas. “Moral Conflict and Political Legitimacy.” *Philosophy & Public Affairs* 16, no 3 (1987): 215–40.
Newbery-Jones, Craig. “Answering the Call of Duty: The Phenomenology of Justice in Twenty-First-Century Video Games.” *Law and Humanities* 9, no 1 (2015): 78–102. https://doi.org/10.1080/17521483.2015.1042218.
Newbery-Jones, Craig. “Ethical Experiments with the D-Pad: Exploring the Potential of Video Games as a Phenomenological Tool for Experiential Legal Education.” *The Law Teacher* 50, no 1 (2016): 61–81. https://doi.org/10.1080/03069400.2016.1146465.
Organisation for Economic Co-operation and Development (OECD). *Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions*. OECD Publishing, September 2025. https://doi.org/10.1787/795de142-en.
Pettit, Philip. *Republicanism: A Theory of Freedom and Government*. Oxford University Press, 1997.
Qin, Hua Xuan, Yuyang Wang and Pan Hui. “Identity, Crimes, and Law Enforcement in the Metaverse.” *Humanities and Social Sciences Communications* 12, no 1 (2025): 1–15. https://doi.org/10.1057/s41599-024-04266-w.
Rawls, John. *A Theory of Justice*. Rev. ed. Harvard University Press, 1999.
Seinfeld, Sofia, Jorge Arroyo-Palacios, Guillermo Iruretagoyena, Ruud Hortensius, Luis E. Zapata, David Borland, Beatrice de Gelder, Mel Slater and Maria V. Sanchez-Vives. “Offenders Become the Victim in Virtual Reality: Impact of Changing Perspective in Domestic Violence.” *Scientific Reports* 8, no 1 (2018): 2692. https://doi.org/10.1038/s41598-018-19987-7.
Shapiro, Andrew L. “The Disappearance of Cyberspace and the Rise of Code.” *Seton Hall Constitutional Law Journal* 8 (1998): 703–23.
Sharma, Vivek. “Introducing a Personal Boundary for Horizon Worlds and Venues.” Meta, February 4, 2022. https://about.fb.com/news/2022/02/personal-boundary-horizon.
Shore, Alexis, Kelsey Prena and James J. Cummings. “To Share or Not to Share: Extending Protection Motivation Theory to Understand Data Sharing with the Police.” *Computers in Human Behavior* 130 (2022): 107188. https://doi.org/10.1016/j.chb.2022.107188.
Slater, Mel, Bernhard Spanlang, Maria V. Sanchez-Vives and Olaf Blanke. “First Person Experience of Body Transfer in Virtual Reality.” *PLoS One* 5, no 5 (2010): e10564. https://doi.org/10.1371/journal.pone.0010564.
Sparrow, Lucy. “How to Govern the Metaverse.” *Wired*, October 19, 2021. https://www.wired.com/story/metaverse-video-games-virtual-reality-ethics-digital-governance.
Suzor, Nicolas P. *Lawless: The Secret Rules That Govern Our Digital Lives*. Cambridge University Press, 2019.

Teubner, Gunther. *Constitutional Fragments: Societal Constitutionalism and Globalization*. Translated by Gareth Norbury. Oxford University Press, 2012.

VRChat Inc. "VRChat Safety and Trust System." https://docs.vrchat.com/docs/vrchat-safety-and-trust-system. Accessed October 4, 2026.

Wang, Na, Jin Zhou, Jie Li, Bo Han, Fei Li and Songqing Chen. "HardenVR: Harassment Detection in Social Virtual Reality." In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 94–104. IEEE, 2024. https://doi.org/10.1109/VR58804.2024.00033.

Werbach, Kevin. *The Stablecoin Toolkit: Financial and Market Dimensions*. Wharton Blockchain and Digital Asset Project (BDAP), The Wharton School, University of Pennsylvania, 2026. https://bdap.wharton.upenn.edu/wp-content/uploads/2026/01/Stablecoin-Toolkit.pdf.