Toward a Constitution for the Metaverse:
Key Principles for the Creation Process

One united humanity.
One shared ecosystem.
Stronger together.

A think-piece authored by
Emily E. Arnold-Fernández
for **OneShared.World**

With contributions from
Jamie Metzl
Shagun Sethi
Sandi Gendi
James Clarke-Lister

Toward a Constitution for the Metaverse:
Key Principles for the Creation Process

Table of Contents

Introduction **2**

Dangers of Lawless Digital Space **3**

A Wide Array of Dangers 3

Digital Harms Often Have Dire and Widespread Impact 4

Many Digital Harms Remain Largely Overlooked 4

A Metaverse Constitution Can and Should Address The Varied Forms of Digital Harm 5

Prior Efforts to Uphold Rights **6**

A History of Permissiveness 6

Corporate Assumption of Rulemaking 7

Civil Society Promotion of Human Rights Standards 9

Governments and Multilaterals Reclaim Internet Rulemaking 13

United Nations 13

States 15

Beyond the Metaverse: Constitution Creation Processes in the Physical World 16

A Way Forward: Rights in the Metaverse **17**

An Inclusive Process is a Moral Imperative and Practical Necessity 17

Recommendations for an Inclusive Process 18

1. Enlist and Empower Champions 18

2. Communicate Proactively and Accessibly 19

3. Reach Constituents Where They Are 19

4. Consider Distributed Volunteer Models to Analyze Input 19

A Constitution for the Metaverse: The Time Is Now 20

Introduction

Digital worlds today are often largely lawless spaces. Where rules are enforced, they tend to be either feudal, imposed by the lord of a particular part of the digital world (such as a specific platform) and subject to few of the ordinary legislative or interpretative processes of a more democratic legal system; or cultural, relying on a shared sense of etiquette that evolves rapidly and encouraged (but not enforced) by social mechanisms. Neither of these approaches leaves room for the safeguarding of fundamental rights or the equitable inclusion of those with less power.

As humans increasingly live significant portions of our lives online – in the metaverse rather than the physical universe, or in a physical universe enhanced by virtual additions to our reality – the dangers of digital lawlessness or, conversely, undemocratic rulemaking and enforcement, pose deeper and wider risks to our mental, physical, emotional and spiritual wellbeing.

To counter this, we argue there is a deep need to create a "Constitution for the Metaverse" with broad, well-designed worldwide public and multistakeholder participation. While the creation of such a Constitution can and should take account of state-led multilateral processes such as the Global Digital Compact, such efforts are inherently exclusive and thus risk being seen as illegitimate or, in more practical terms, meeting widespread resistance. Civil society efforts such as the African Declaration on Internet Rights and Freedoms provide intriguing alternatives but so far remain regional rather than global; they also may lack necessary buy-in from more powerful actors such as governments and the private sector. We believe a broader approach is needed: one that centers current and future denizens of the metaverse to create a constitution that reflects the values by which they wish to be governed.

Responding to the most immediate gap in current governance of digital spaces, we propose the central component of this Constitution should be a Bill of Rights or other definition of the fundamental rights the Constitution is designed to safeguard. To ensure this Constitution and Bill of Rights has

broad legitimacy, we believe this process should be informed by principles of democratic and inclusive norm-making and rulemaking. This memo attempts to summarize the history of efforts to date, identify key gaps in processes and outputs so far, and propose the foundational principles of a more inclusive and legitimate process to create a true Constitution for the Metaverse.

Dangers of Lawless Digital Space

A Wide Array of Dangers

The dangers in digital space are, in broad terms, not significantly different from those in physical space. Individuals can become the targets of malicious speech intended to directly impair their mental wellbeing, or to convince them to engage in activity that directly harms their physical wellbeing; or negligent speech that has similar effects, albeit not directly intended. Individuals may have their digital property stolen, including money – not just cryptocurrency, but also digital proxies for physical cash such as bank balances. They may experience identity theft. They may be coerced into consenting to harmful or dangerous practices through the use of contracts of adhesion, where such contracts serve as a gateway to accessing resources or community.

Beyond harms that occur more or less entirely in a digital space, there are those where digital and physical worlds interact in ways that enable or magnify harm. Speech online can be tracked to physical locations, such that governments or private actors can engage in violence against those engaged in certain types of speech, or who possess characteristics such as a particular ethnicity or gender. Digital spaces can be used not only for malicious speech intended to harm listeners, but also for speech intended and even deliberately designed to incite listeners to harm others. Virtual worlds provide venues for conspiracy to commit all manner of harms, from human trafficking and enslavement to un- or under-compensated extraction of resources from places or people with limited power to resist.

Digital Harms Often Have Dire and Widespread Impact

While these harms may seem abstract, their impact can be severe. The Myanmar military government has detained or killed activists after tracking their online speech. A white supremacist radicalized online murdered Black shoppers in Buffalo, New York. Teens have been encouraged to plan and commit suicide in online forums – and some have done so. Game developers Zoë Quinn and Brianna Wu, among others, received online threats of rape and murder; Wu has publicly shared that she was diagnosed with post-traumatic stress disorder as a result.

These harms also can directly impact large numbers of people. At any given moment, over 40 million people are experiencing labor or sex trafficking or forced marriage, with women and girls accounting for 71 percent of this figure. Traffickers are increasingly using digital spaces to lure people into trafficking situations. Perhaps less imminently drastic but with the potential to grow ever more problematic, digital colonialism through the extraction and ownership of data with minimal benefit to the people who are sources of that data, and a deepening divide between the privileged and those who are not, also affect millions or billions around the world.

Many Digital Harms Remain Largely Overlooked

Much of the scholarship on digital harms has, until now, focused on government limitation or suppression of free speech online. Such harms can be real: The governments of China and Myanmar, among various others, have used physical violence and incarceration to silence and at times kill critics.

Many of those writing about the dangers of government suppression of free speech, however, have failed to take into account other dangers, for example hate speech and incitement that can also foment violent or otherwise harmful action, or digital colonization, perhaps in part because the writers occupy privileged positions in the Global North in which they are unlikely to be direct targets of harms such as misogynistic or racist violence or the corrosive impacts of colonization.

Moreover, while digital worlds are, today, still largely mediated by tools that humans can set aside, the harms that occur often transcend those tools, encroaching into the physical world. Where the harms occur purely in digital spaces, humans may still have very real experiences of mental or emotional harm; and where humans already interact with concepts that have an imperfect physical analog, such as money, digital attacks can result in analogous physical-world impacts. As digital and physical worlds grow ever more integrated, as some believe likely, the types of harms that occur at least partly in digital worlds is likely to grow.

A Metaverse Constitution Can and Should Address The Varied Forms of Digital Harm

To be clear, it is not necessary to compare or rank different forms of digital harm in order to devise a Constitution or Bill of Rights for digital worlds. Humans hold a range of rights that are fundamental to a full and thriving human existence, online no less than in the physical world. Rather, the creation of a Constitution requires consideration of the full scope of rights that are and will be fundamental to human wellbeing, and ultimately the codification of these rights into a governing instrument that can serve as a social compact between all individuals and entities acting in the digital worlds to which we collectively refer as the Metaverse.

As we consider an appropriate Constitution for the Metaverse, it will be critical to consider the wide range of harms that may affect people in disparate circumstances and physical locations and with disparate characteristics. It also will be important to attempt to envision harms that may emerge or become more prevalent over time. Finally, it will be important to develop a fair and effective process for continued adjustment of the Constitution based on changes we, as yet, cannot imagine.

Prior Efforts to Uphold Rights

A History of Permissiveness

In digital spaces, relatively little has been done to comprehensively define or safeguard rights. Although efforts in the 2010s and 2020s have begun to catalyze discussion about human rights protections in the metaverse, the early history of digital spaces was one of non-regulation and self-organization.

This permissiveness in part resulted from a widespread belief, during the emergence of the internet as a popular medium in the US in the 1990s and early 2000s, that governments posed the greatest systemic risk to individual rights, particularly the civil and political rights that were understood as most relevant to digital worlds at the time. As a consequence, prominent US public voices advocated for limited to no government regulation of digital spaces.

The most libertarian of these voices were activists such as Grateful Dead lyricist John Perry Barlow and technology entrepreneurs Mitch Kapor and John Gilmore, who argued fiercely against virtually any form of regulation or structured norm-setting. These activists – who went on to found leading internet policy advocacy organization Electronic Frontier Foundation (EFF) – saw governments as threats to, rather than protectors of, human wellbeing in the digital world. Their views were formed, in part, by experiences with US government officials attempting to apply existing laws or create new ones for digital spaces based on insufficient technical knowledge; these attempts threatened or resulted in outcomes that were at odds with the activists' cultural norms about justice.[1]

A more tempered view was advanced by advocates such as Lawrence Lessig, a prominent lawyer and scholar (also at one time a board member of EFF),

[1] Interestingly, although these activists and EFF as an entity advocated against governmental regulation of digital spaces, their objections were still couched in the language of law. In a ringing declaration delivered at Davos in 1996, Barlow proclaimed, "we do not need a constitution or government in cyberspace, all are free and independent" – yet in his own history of EFF, he writes that EFF was created because he and other founders objected to violations of the US Bill of Rights and felt the need to defend the US Constitution.

who acknowledged a need for codification and safeguarding of rights while also expressing concern about the contemporary implementation of government regulation of digital spaces. Lessig foresaw a future in which governments would require the architecture of digital spaces (e.g., software code) to enable instant, continual identification of all participating individuals, such that anyone existing in digital spaces could be tracked or blocked. Lessig's vision, however, failed to foresee that instant, continual identification would become the norm in digital spaces not because a government required it, but because the creation of digital spaces was left to private for-profit corporations who benefited from such architecture.

Views such as these contributed to a permissive tendency toward digital spaces by governments in the 1990s and early 2000s. Digital worlds grew in scope, but remained largely unregulated. Even where governments took an authoritarian approach to regulating digital spaces, this approach was often narrow in focus: In the Global South / East, governments tended to impose rules related to critique of their regimes, while in the Global North / West, governments tended to impose rules designed to suppress certain narrow types of speech seen as fundamentally undermining the morality of the country (for example, in the US, obscene or pornographic speech accessible to minors; in Germany, hate speech that extolled Nazi ideology).

Corporate Assumption of Rulemaking

While governments shied away from broad regulation of digital spaces, other entities took control of digital spaces by virtue of their ownership of the property and infrastructure necessary for creation of these spaces. Most of these entities are for-profit corporations, often very large and, by virtue of their size, with significant influence in the global economy. As their power has grown, these companies have also come to hold significant influence on society, culture, and politics. Some of the most influential include Google, Facebook, and Apple among others.

While rulemaking in the early days of the digital world resulted primarily, as Lessig notes, from the architecture of the code (and hardware) that enabled these spaces, the corporations that came to largely own digital space not only

adapted its architecture to serve their own interests, but also gradually imposed other usage rules. Today, usage rules range from anti-piracy agreements to prohibitions on hate speech; architecture rules often include tracking of movement and activity, and submitting to propaganda (such as advertising) as a condition of entry.

Architectural rules are enforced automatically, typically with no opportunity to negotiate or appeal to a decisionmaker capable of considering questions of justice. Enforcement of usage rules happens both directly, through corporate acts such as removing or limiting a violator's access to a digital space, and mediated by traditional state-run enforcement mechanisms such as police and courts; in both cases, appeal of an initial enforcement action is possible, albeit sometimes inaccessible to those without certain privileges. Both architectural and usage rules vary by corporation, such that different sets of rules may be enforced in different "territories" of the digital world. Unlike physical territory, however, a person may be active in multiple territories at once, for example when reading a New York Times news article within Google's Chrome browser, or scrolling through Facebook on an Apple device.

Because much enforcement of corporate rulemaking in digital spaces is conducted either by the infrastructure itself or by the corporation, such enforcement is designed to serve the interests of corporate owners. Indeed, many would argue that corporations, at least those that are publicly traded (as the largest digital space owners are) have a legal obligation to ensure their rulemaking and enforcement systems serve the corporation first and foremost, as part of every corporation's purported duty to maximize profits for its shareholders. In the physical world, corporate ownership of large amounts of territory is widely seen as problematic precisely because of a virtually universal philosophy that territory should be governed for the good of those who inhabit it[2] rather than exclusively for the benefit of those who own it. In

[2] The British East India Company's legacy in India is an illustration of some reasons such ownership is problematic. By contrast, both democratic models such as that of Canada or Kenya and communist models such as China or Morocco are founded on the ostensible premise that the purpose of governance is the good of citizens (even where the practice of governance patently gives the lie to this idea).

the digital world, however, corporate ownership is currently expected and seen as normal.

Although corporate rulemaking for digital worlds is generally designed to benefit the corporation, calls for corporate adherence to human rights norms (from both inside and outside corporations that control digital spaces) have borne some fruit. Google, for example, has attracted talent since its inception by publicly proclaiming its "Don't be evil" motto. More recently, Facebook has created an advisory board to adjudicate user appeals of Facebook content moderation decisions. The advisory board's charter and bylaws, among other documents, indicate that Facebook expects the board's decisions to be guided by international human rights norms and principles, at least those "protecting freedom of expression". Some commentators argue that corporate business interests, in particular an attempt to stave off government regulation, ultimately drive such initiatives. Nonetheless, initiatives such as these may offer ideas for developing the design process and the content of a Constitution for the Metaverse.

Civil Society Promotion of Human Rights Standards

Alongside corporate efforts to grapple with human rights norms in the digital world, independent civil society initiatives also have made strides toward creating human rights rules and norms for digital spaces. Civil society efforts have grown as the dangers of government permissiveness and corporate control have become increasingly apparent with the expansion of the metaverse. As David Souter, lead consultant on the ten-year review of the United Nations' World Summit on the Information Society, has noted, "It's generally recognized that the 'wild west' days of early internet are over and that regulation's going to be fundamental to IG [internet governance] in the future." Civil society organizations concerned about human rights in the digital world are increasingly engaged in research, analysis and advocacy to define and promote adoption of human rights standards for the metaverse.

In 2014, Human Rights Watch called for the development of governance structures to protect human rights in digital spaces: a United Nations Special Rapporteur, a stronger multistakeholder internet governance model, and a

conceptual reframing of human rights as essential, rather than antithetical, to national security in digital spaces. These recommendations, however, did not seek to define human rights in the metaverse, or to apply existing substantive human rights standards to digital spaces. More recently, two significant civil society initiatives to propose and promote human rights standards in the metaverse have emerged: The African Declaration on Internet Rights and Freedoms (AfDec) and Ranking Digital Rights' *Big Tech Scorecard*.

Created by a coalition of mostly Africa-based civil society organizations, the African Declaration articulates thirteen principles for upholding human and people's rights in the digital world and describes how these principles should be applied, with a particular focus on "Africa's social and economic development needs and goals". The Declaration builds on a human rights legal instrument initially developed to govern physical spaces, the African Charter of Human and People's Rights, as well as civil society standard-setting initiatives such as the Windhoek Declaration and the African Platform on Access to Information.

AfDec's principles are:

1. Openness
2. Internet Access and Affordability
3. Freedom of Expression
4. Right to Information
5. Freedom of Assembly and Association and the Internet
6. Cultural and Linguistic Diversity
7. Right to Development and Access to Knowledge
8. Privacy and Personal Data Protection
9. Security, Stability and Resilience of the Internet
10. Marginalized Groups and Groups at Risk
11. Right to Due Process
12. Democratic Multistakeholder Internet Governance
13. Gender Equality

While AfDec's principles are more comprehensive than other attempts to define human rights in digital spaces, it is not always clear whether these principles apply only to government regulation, or to the behavior of other

actors such as private corporations. For example, Principle 5, Freedom of Assembly and Association, says in part:

> Everyone has the right to use the Internet and digital technologies in relation to freedom of assembly and association, including through social networks and platforms. *No restrictions on usage of and access to the Internet and digital technologies in relation to the right to freedom of assembly and association may be imposed* unless the restriction is prescribed by law, pursues a legitimate aim as expressly listed under international human rights law (as specified in Principle 3 of this Declaration) and is necessary and proportionate in pursuance of a legitimate aim....Everyone should enjoy unrestricted access to the Internet. *Any shutting down or blocking of access to social networking platforms*, and in fact the Internet in general, constitutes a direct interference with this right (emphasis added).

While nothing in the description of this principle and its application explicitly states that it applies only to states, the description also does not appear to consider limitations imposed by non-state actors such as the corporations that control social networks. The language used in the AfDec would not allow Facebook to block users spreading misinformation about COVID-19 (because such removal is not prescribed by law) or Twitter to ban former US President Donald Trump under its prohibition on "glorification of violence" (because Twitter's policy bans speech that is well outside the legal definition of incitement).

To be sure, AfDec's discussion of Principle 3, Freedom of Expression, does address the case of non-governmental actors. AfDec calls on "intermediaries" that "operate self-regulatory systems and/or make judgment calls on content and privacy issues" to do so in ways that align with human rights standards, although it still makes no recommendation regarding external accountability. While AfDec's signatories might argue that government suppression of free expression is a more immediate threat given the prevalence of internet shutdowns and government ownership or licensing of internet providers, AfDec's failure to set forth clear human rights obligations for private entities limits its direct use in regulating non-governmental conduct in the

metaverse. Nonetheless, the human rights principles it articulates could be a starting point for discussions of a broader Metaverse Constitution that governs the conduct of all actors in the digital world.

US nonprofit think tank New America's independent research project, Ranking Digital Rights, "evaluates and ranks the world's most powerful tech and telecom companies on their commitments to respect users' fundamental rights, and on the mechanisms they have in place to ensure those promises are kept," reporting its findings on the *Big Tech Scorecard*. RDR's methodology uses the Universal Declaration of Human Rights as its foundation, specifically focusing on rights related to governance, freedom of expression, and privacy. Much of RDR's emphasis is on corporate transparency about their internal rules and processes; responsibility for preventing or redressing harms is largely missing from RDR's scoring methodology.

While RDR's scorecard is a step toward incentivizing conformity with a narrow band of human rights procedural standards, it does little to address corporate responsibility for substantive human rights implementation or enforcement in the digital spaces under corporate control. In addition, nothing in RDR's approach addresses corporate responsibility for safeguarding rights outside of those directly related to free expression and privacy within corporate-controlled digital territory.

For example, while a corporation may receive a lower score if it is not transparent about when and how it will share user data with a government, a corporation that follows its transparent rules about data-sharing in a way that leads or contributes to an activist's death will not be penalized in the ranking. Moreover, a corporation that controls a digital space where one person stalks or harasses another will not be penalized in the ranking, even if the corporation did not abide by its own enforcement policies in permitting the stalking or harassment to continue, provided it was transparent about prohibited activity and included and advertised features to allow the user to keep their account secure. By analogy, if RDR were ranking governments, a government that published its laws, and provided residents with free door

locks and installation instructions, would have no further responsibility per the RDR ranking for preventing or redressing crimes.

Governments and Multilaterals Reclaim Internet Rulemaking

As corporations increasingly come under fire for human rights abuses perpetrated partly in digital spaces, and civil society calls for digital human rights protections grow more specific, governments and multilateral institutions have sought to reclaim and expand their rulemaking power vis-a-vis digital spaces.

1. United Nations

The United Nations is slowly progressing toward the creation of a Global Digital Compact – which could ultimately serve as a Constitution of the Metaverse. The Compact, which aims to define "shared principles for an open, free, and secure digital future for all," is a component of the UN Secretary General's Roadmap for Digital Cooperation, overseen by his Envoy on Technology, a position created in 2019. Initial development of the Global Digital Compact is happening via a process of public, multistakeholder contribution and review, with a final version to be negotiated at the proposed 2023 Summit of the Future.

While the Global Digital Compact process currently appears to align with best practices in inclusive rulemaking and norm development for digital spaces, there are reasons to be skeptical as well. The rise of non-binding compacts over the past decade as a substitute for binding international legal instruments addressing human rights arguably is problematic, as it weakens the accountability of states and other actors by replacing rules with suggestions. Where compacts do contribute to norms, they may encourage eventual legal recognition of new rights – but they may also undermine existing legal rights by encouraging a narrower or more limited interpretation of binding human rights instruments.

Additionally, while the Global Digital Compact process currently invites open contributions from all people, the process does not appear widely advertised at time of writing. The Compact appears to have no coverage on Reddit, a

major internet discussion forum, nor in the largest international wire services: Reuters and Associated Press. The Compact also apparently has not received any coverage in the New York Times, the Guardian, or the Washington Post. Single brief mentions of the Compact have appeared in other newspapers including the Guardian Nigeria (not affiliated with the Guardian), El Mostrador in Chile, Straits Times, and Foreign Policy in the US, but by any measure the Compact is not being widely publicly discussed, two months after its concept note and contribution portal were publicly shared by the UN. Coupled with the reality that states still control virtually all United Nations processes, this suggests the final drafting and decisions about the Compact's content may ultimately reflect a more Westphalian than multistakeholder approach.

In addition to the Global Digital Compact and the Summit of the Future, the United Nations also has other avenues for influencing rulemaking for digital spaces. One is the Internet Governance Forum (IGF), created in 2006 with a mandate to facilitate discussion of digital world policies and best practices. IGF could play a leadership role in convening relevant stakeholders to create a Constitution of the Metaverse, although as an institution created by UN mandate, it – like the Global Digital Compact – might tend toward a state-centric, Westphalian approach rather than a truly multistakeholder agreement.

Another United Nations entity addressing digital human rights is the Special Rapporteur on the Right to Privacy, established in 2015.[3] The Special Rapporteur on Privacy is charged with advocating and supporting government changes to bring surveillance in line with rule of law and human rights, and laying out the obligations of the private sector with regard to human rights protections. A Special Rapporteur, however, typically reports on and advises specific countries or regions, rather than prescribing global norms that could serve as a foundation for a Metaverse Constitution. Where a Special Rapporteur does make broader global recommendations, these

[3] Civil society has largely sought and welcomed increased multilateral regulation of digital spaces. For example, Human Rights Watch's call for a United Nations Special Rapporteur on the Right to Privacy became a 90-organization advocacy effort that resulted in the appointment of the special rapporteur in 2015.

usually focus on a narrow thematic issue (for example, privacy protections during a pandemic).

2. States

In addition to United Nations efforts, states are increasingly seeking to expand their rulemaking in digital spaces, including to prevent and address harms against people. While rule enforcement generally takes place at the national level, many of these efforts have included a multilateral component, albeit usually outside of formal multilateral structures such as the United Nations. These expansions arguably began in the 2010s with concerns about privacy (as Lessig's predictions about the tracking implications of digital infrastructure design were realized in corporate-controlled spaces), but recent efforts are far more comprehensive.

In Europe, for example, the General Data Protection Regulation went into effect in 2016 as an effort to protect individuals' data and privacy in digital spaces, changing the behavior of virtually all who control licit digital spaces open to the public.[4] A few years earlier, in 2011, thirteen countries established the Freedom Online Coalition, focused on core civil and political rights such as free expression, association, and assembly, as well as privacy.

Today, however, state ambitions are far broader: In April 2022, 60 governments signed a Declaration for the Future of the Internet, a US government-led attempt to establish broad principles applicable to all digital worlds. This Declaration gets closer to articulating a comprehensive set of ideas that might serve as a foundation for a Constitution of the Metaverse. The core principles of the Declaration include:

- Protection of Human Rights and Fundamental Freedoms
- A Global Internet that operates on principles of net neutrality and free flow of data, without government-imposed shutdowns or degradation
- Inclusive and Affordable Access to the Internet

[4] Another example is the 2018 California Consumer Privacy Act. Although only binding at the sub-national level and only from 2020 onwards, it also has impacted the behavior of those who control digital spaces.

- Trust in the Digital Ecosystem, including protection of privacy, human rights, electoral processes and online security while combating unfair, malicious, or criminal online activity
- Multistakeholder Internet Governance

The Declaration also recognizes United Nations and other multilateral efforts and positions itself as complementary.

Two characteristics of the Declaration's creation process undermine its potential as a basis for a Constitution of the Metaverse, however. First, the process was entirely state-led, rather than multistakeholder, with no direct public input. Those most likely to suffer internet harms were represented only by their governments, if at all – and many of those who are at risk of internet harms live under government regimes that did not participate in the Declaration. Moreover, the Declaration implicitly focuses on harms most likely to be inflicted by governments, without seriously addressing corporate complicity in misinformation campaigns and other harms impacting individuals.

Second, the Declaration originated as an attempt by the United States government to form an alliance of states against China and Russia. A leaked early draft showed a focus on technical cooperation in support of US political and economic interests, with human rights mentioned only once, prefatorily. China and Russia are expected to ignore it, which will undermine claims to global legitimacy. If Chinese and Russian citizens and civil society organizations had contributed to its development alongside other non-state stakeholders, a case for its legitimacy would be stronger – but under the circumstances, it is likely to been seen primarily as geopolitical maneuvering.

Beyond the Metaverse: Constitution Creation Processes in the Physical World

To identify the principles that give legitimacy to constitutional development processes, we need not restrict our gaze to prior efforts in digital spaces. Humans have been developing democratic governance systems for millenia. We can look to recent constitutional processes in the physical world – to

broaden our understanding of how we might enhance the legitimacy of a process for creating a Constitution of the Metaverse.

One recent good practice example is the Kenyan constitutional revision process in 2008-2010. This process, which substantially overhauled Kenya's much-amended post-colonial constitution, combined broad public consultation with expert analysis and development of specific constitutional provisions: Widespread advertisements invited the public to submit their views via memoranda, which an Expert Committee analyzed to understand the strength of popular opinions and the breadth of divergence among them. Subsequently, thematic and sectoral consultations allowed the Experts to refine their understanding of public views. Where needed, the Experts also engaged technical consultants.

The Experts then made a draft constitution available for public review and parliamentary consultation. Based on input from both processes, the experts issued a further draft, to which the National Assembly could propose amendments. Following that process, a final draft of the proposed constitution was subject to public referendum and, when passed, became Kenya's new governing instrument – which remains substantially in place today.

While other modern constitutional processes – for example, that of Ecuador in 2008 or the drafting of the Universal Declaration of Human Rights in the 1940s – each have their own unique details, common elements tend to include public consultation coupled with expert leadership and drafting. The selection of expert representatives with appropriate credentials or experience usually occurs through a process that is essentially democratic or republican in nature; and the selected representatives undertake the responsibilities of conducting public consultation and analysis followed by debating, negotiating and proposing a shared set of rules that is then ratified either by the public or their elected representatives.

A Way Forward: Rights in the Metaverse

An Inclusive Process is a Moral Imperative and Practical Necessity

To effectively uphold human rights in the metaverse, we must create a governing instrument that has widespread legitimacy and buy-in. This is true for any constitution, but it has become particularly critical for digital worlds because of their dualistic history: State permissiveness and early adopter resistance to regulation that has led to lawlessness in many corners of the metaverse, alongside feudalistic, authoritarian regimes of digital landlords who are unaccountable to their denizens. Already governance of the metaverse suffers from a legitimacy deficit; to reverse this, any global governance effort must align consistently with best practices in participatory co-creation.

At the same time, the momentum of current efforts should not be discounted. In particular, the Global Digital Compact has the imprimatur of the United Nations, which remains – despite critiques and flaws – the multilateral governing entity with the greatest global legitimacy. The Compact's biggest drawback is the insufficiency of its public and stakeholder consultation process so far; if this could be remedied, we could harness rather than compete with the energy being invested into the Compact process.

Recommendations for an Inclusive Process

Drawing on strong examples like Kenya's 2008-2010 constitutional revision and the African Declaration (AfDec) development process, the Global Digital Compact process – or another global effort to create a Constitution for the Metaverse – could uphold principles of inclusion and participatory governance by taking the following four steps:

1. Enlist and Empower Champions

A public consultation process is only effective if it engenders ample and diverse participation. To achieve this, the Global Digital Compact process (or other digital constitution processes) should ensure their public consultations are championed by **diverse civil society** that is **well-connected in digital**

worlds or to current or prospective digital denizens; and **willing *and resourced*** (with adequate financial and technical support) to conduct widespread **outreach** and to **support underrepresented populations** to organize and participate.

2. Communicate Proactively and Accessibly

Communication must reach the eyes and ears of its intended audiences. While the Envoy on Technology has launched a relatively user-friendly website for participation, it should also invest in designing and executing a well-constructed **public communications plan** that invites participants – in **multiple languages, media,** and **disability-friendly formats** – to join the Global Digital Compact consultation process. The communications plan should be designed both to **motivate** participation and to provide **clear, simple instructions** on how to do so.

3. Reach Constituents Where They Are

To engender ample and diverse public participation, the Global Digital Compact process (or other digital constitution processes) should disseminate information via a plethora of communication channels in both virtual and physical spaces. These might include: Radio and billboard advertisements in places where the internet is not easily accessible and commonly used; printed and online news media including both traditional media and widely-followed "new" media such as blogs; and social media including Facebook, Instagram, Twitter, TikTok, WhatsApp, WeChat, YouTube, Reddit and the like.

4. Consider Distributed Volunteer Models to Analyze Input

Wide public participation on a global scale poses a challenge, as the scope of review and analysis needed to appropriately consider all input will be extensive. Using a distributed volunteer model, like that used to edit Wikipedia, could make this challenge manageable while increasing the participatory nature of the process. Because multiple volunteer reviewers could contribute to the analysis of any input, the aggregate analysis is likely to be strong. One consideration is ensuring a representative diversity of distributed volunteer reviewers; this has proven challenging for Wikipedia,

but could be addressed through a campaign to encourage volunteers from a diverse range of backgrounds, locations and languages.

A Constitution for the Metaverse: The Time Is Now

The convergence of state, multilateral and civil society efforts to articulate and promote human rights principles for the metaverse over the past decade indicates rising public concern about the current gaps in governance of digital worlds. Much of the world are digital denizens – and we are experiencing in real time the harms that arise in undemocratic online spaces. As more and more of us suffer these harms or know someone who has, a growing body of people across the world is calling for a new human rights instrument: A Constitution for the Metaverse.

Although the metaverse differs from the physical world, its inhabitants are the same: us. The time is now to safeguard our human rights where we spend increasing amounts of our lives – in the metaverse.