1

Introduction

Comparative, Generative, and Synthetic Analysis of International Dimensions of Platform Responsibility

Bhaskar Chakravorti and Joel P. Trachtman

From the giants – Facebook, YouTube, TikTok, and WhatsApp – to the corners occupied by Discord, Caffeine, and NextDoor, there is no scarcity of digital platforms to browse and in which to lose onself.[1] Social media has come a long way since its early days of Six Degrees or Friendster. More than half the world's population is on one or more such platforms and they spend 15 percent of their waking hours on them. Collectively, they watch a billion hours of YouTube videos a day and between sharing personal updates, pictures of the last meal, breaking news, and outright falsehoods, half a billion stories are posted on Facebook every day.[2] It is no wonder, therefore, that social media platforms have enormous power to shape narratives and influence behavior.

This raises the question of who bears responsibility for what appears on the platforms, how the content is shared, and who gets to see it. Users cannot be expected to – reliably and consistently – self-regulate their posts. It is natural to expect that in the absence of a set of rules and editorial intervention, the communicative power and global nature of the platform space will be severely compromised. The responsibility of the platforms for the use of their power to control and disseminate content is statutorily limited by national laws, such as Section 230 of the Communications Decency Act in the US or the Digital Services Act in the EU. Legislative efforts to demand and guide appropriate content moderation and to avoid private abuse of this power offer essential safeguards, but they are in tension with several competing concerns. Any restraints on content must be balanced against the commitment in liberal states to avoid excessive government regulation, especially of speech. Less liberal states will have different policies, and even liberal states differ with respect to their priorities. The US focuses on free speech,[3] while the EU more frequently compromises free speech for other values. Second, users must

[1] The idea of a "platform" itself merits some discussion. We shall defer it to a later chapter by Eric Goldman (Chapter 2).

[2] *Top social media fun facts in 2023*, Amra & Elma (2023), https://www.amraandelma.com/social-media-fun-facts/, (last visited Mar. 2, 2024).

[3] *See* Michael Glennon, Free Speech and Turbulent Freedom (2024).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001

have the opportunity to exercise their creative urges and possibly develop businesses based on the content they post. The tools to exercise such creative urges have been given an added boost by the wide availability of applications enabled by generative artificial intelligence (AI). Third, social media offers a powerful way to tap into the "wisdom of the crowd" for information and bottom-up insights on a wide range of issues from breaking news stories to reviews of products and any restraints ought not to skew the outcomes.

Yet another complication is that digital platforms are borderless (except in countries where they are blocked), while state-imposed rules and regulations that govern content on platforms are country-specific. Diverse – and sometimes contradictory – national rules can splinter platforms and potentially undermine their utility. Intermingled among all of these issues are the very different sets of political, commercial, and personal interests of the principal actors – users, platforms, and regulators. As we look ahead, a fourth actor is poised to become a principal as well: AI that may not have "interests" that drive the decisions made by underlying algorithms but is an autonomous player that can evolve beyond the control of its human creators and their intentions.

With these issues as context, can a global approach be developed to address them while elevating the quality of content and limiting material that could lead to harm? This is the question that motivated the book.

As we consider setting boundaries around harmful content, there are four challenges to be addressed:

- *Volume*. First, platforms host primarily user-generated content; the number of users and the amount of content are enormous, as is the frequency of posting – billions each day. Platforms make it costless for many people to share content among varying and potentially very large groups. Where user-generated content may violate national substantive laws, such as those applying to defamation, pornography, intellectual property protection, consumer protection, trafficking, pornography, election protection, etc., it is difficult for national regulators to keep up. While it is easy to say that what is illegal in the real world is also illegal in the virtual world, the virtual world raises novel problems of enforcement. Some states respond by delegating some aspects of enforcement to platforms themselves, sometimes leaving it to the platforms to establish private regulation that may exceed the powers of government regulation, especially in free speech areas.
- *Diversity*. Second, regulators and lawmakers in every country are guided by their own legislative, bureaucratic, and political agendas. Each might operate at a different pace, and each might want to push forward regulations to suit local conditions and political and economic interests. As a result, while reform of Section 230, the internet platforms' current

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001

intermediary liability regime in the US, is being debated, other jurisdictions, notably, India, Brazil, Australia, Kenya, the UK, and the EU, have taken or are considering taking their own actions to hold platforms responsible for content. This could result in a chaotic patchwork of rules worldwide. The patchwork of rules could splinter the platforms in ways that are costly not just to the platforms, but to society at large.

- *Indivisibility*. Third, to a significant extent, platforms are valuable precisely because they are global, establishing a global common communications space and providing generous economies of scale and scope. This very global nature makes platforms vulnerable to divergent national regulation. States set rules focused on their own territory, but these rules inevitably have effects beyond their territories. On the other side, states that wish to have different rules – to assert digital sovereignty" – may be unable to effect their policies, especially if they are not the home state of the platform. While technological devices can be used to reestablish "borders" for platforms, these may be incompletely effective.
- *Scarcity*. Finally, there could be unintended consequences of well-intentioned actions by regulators and lawmakers of any one country; such unilateral action might skew the incentives for the social media companies in ways that could, ironically, make matters worse globally. The platforms could respond to regulations imposed by one country, say the US, by over-allocating resources to ensure protections for users in the US and not run afoul of the laws; but given limited resources, this could come at the cost of resources under-allocated elsewhere. As an example, consider the 2021 Facebook papers revelations that the company had allocated 87 percent of its content moderation resources to monitoring US content, even though less than a tenth of all Facebook users are in the US. Moreover, the greatest risks of dangerous content are in societies with weaker political and legal systems and with more vulnerable populations. Aggressive action by US regulators could result in outcomes that make conditions worse for these societies unless the new regulations proactively create mechanisms to guard against such risks. The worst-case scenario is that societies with the least resources and low global "share of voice" suffer the most as the companies must respond to other demands.

If the overarching goal is to "defeat" disinformation – a term we use to cover a broad range of harmful content – the failure to address these discontinuities will result in potentially insufficient and inefficient restrictions.

In response, The Fletcher School carried out a timely and novel collaboration between two key centers, its Institute for Business in the Global Context and its Center for International Law and Governance, to coordinate expert analysis and

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001

research designed to examine the dynamics of these challenges and their possible resolutions. We began with the specific question of reforming rules that make platforms responsible or immunize them from responsibility for content they host, ranging from modification of Section 230 to more focused rules requiring moderation and takedown, defining hate speech, restricting non-consensual image-sharing, defining cyber bullying, prohibiting misinformation to consumers, and outlawing threats to elections and civil discourse, etc.

As we investigated the question of platform responsibility, we examined whether there is useful guidance or precedent offered by other regulations and widely agreed-upon international standards in other fields, such as international finance, international taxation, international public health, and initiatives aimed at countering violent extremism, that govern companies and have cross-border implications.

We assembled a set of experts to contribute chapters aimed at a broad audience of policymakers, lawmakers, and industry decision-makers. The chapters are organized in three categories: comparative, generative, and disciplinary analyses.

Comparative Chapters. The comparative chapters examine specific countries' policies, and the bases for those policies, in this area, seeking to understand how different national social goals and structures demand different types of rules. These chapters include leading technologically advanced jurisdictions like the US, the EU, China, India, and Brazil. These chapters were prepared by experts in the relevant area and contain in-depth analyses responding to specified questions and based on experience, study, and interviews with leading thinkers and policymakers. Eric Goldman (Chapter 2) summarizes the legal framework governing social media platforms in the US by highlighting three pillars: the constitutional protections for free speech and press, Section 230, and the limits on state regulation of digital platforms. Christoph Busch (Chapter 3) provides an overview of the EU regulatory framework and its nuanced "due diligence" approach to content moderation, along with the potential international effects of EU rules on platform responsibility. Jufang Wang (Chapter 4) considers the general principles that govern China's online content governance and how it exercises governmental control of its information space, along with a brief case study on TikTok. Artur Pericles (Chapter 6) characterizes Brazil as a "battleground" where proposals for platform responsibility have been advanced and disputed; he provides an overview of existing and proposed frameworks, and of recent developments there and the lessons they offer. Finally, Jhalak Kakkar, Shashank Mohan, and Vasudev Devadasan (Chapter 5) provide an overview of the regulatory frameworks that govern content on social media platforms in India and their evolution as well as the challenges and petitions that have been filed as the rules have evolved.

The overall picture is one of considerable diversity across countries and steady evolution. The latter is often in response to a consideration of rules in place in other regions or to external events where social media content has played a role or

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001

petitions and challenges to existing rules. This initial set of chapters help with setting the stage for many of the key issues that must be addressed.

Generative Chapters. The generative chapters examine other partially analogous global regulatory issues to generate proposals for addressing the issues at both the rule-making stage and the rule-application stage. Federico Lupo-Pasini (Chapter 8) focuses on the Basel Committee on Bank Regulation capital requirements on banks and the potential lessons from financial regulation. Carlo Garbarino (Chapter 9) considers the OECD Base Erosion and Profit Shifting International Taxation Project. Mark Jit and Dominik Hofstetter (Chapter 7) offer a comparison based on the global response to infectious disease, examining how the international community might respond to "misinformation pandemics." Farah Pandith and Simone Lipkind (Chapter 10) explore the experience of countries looking to counter violent extremism. These examples have a crucial factor in common: they deal with issues that are subject to national regulation and enforcement but have significant implications beyond national boundaries. The generative chapters explain the global scope of the problem, detail the process and institutional structures by which rules have been promulgated and applied, and briefly compare these contexts to and suggest lessons learned for content moderation and platform responsibility rules.

Disciplinary Analyses. The issues at the heart of defeating disinformation on social media platforms are complex and defy straightforward solutions. Clearly, even the insights drawn from a breadth of comparative and generative analyses can fall short of delivering answers. The reason for this is that the underlying factors have disciplinary characteristics that draw upon a combination of microeconomics, international law, international politics, and technology policy. Each disciplinary lens offers insights into the drivers behind possible solutions and the barriers that might get in the way. This set of chapters considers these drivers and barriers and offers potential paths forward in addressing the complexities using the lens of these individual disciplines. Josephine Wolff, adopting a technology policy perspective, considers different policy goals related to social media platforms along three different types of obligations: responsibilities to target particular categories of unwanted content, responsibilities for platforms that wield particularly significant influence, and responsibilities to be transparent about platform decision-making. Wolff (Chapter 11) explores which of the policy goals present the greatest opportunities for international coordination and agreement and the lessons learned from the comparative and generative chapters. Daniel Drezner (Chapter 12) argues, from the perspective of an international political scientist, that we are destined for what he describes as a hypocritical system of "sham governance," with token agreements negotiated at the global level but with inadequate enforcement mechanisms. He concludes that regulation is likely to remain national – and fragmented. Bhaskar Chakravorti (Chapter 13) adopts a microeconomist's perspective and examines the incentives structure that are at the heart of social media platforms run as

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001

businesses and discovers that they lead to a "disinformation paradox." The attempts to moderate content by regulators acting independently across nations leads to an increase in harmful content worldwide. Joel Trachtman (Chapter 14) observes that the approach of international law is to address issues vertically relating to specific regulatory areas, such as human rights, election integrity, consumer protection, privacy, defamation, human trafficking, competition, tax, etc., rather than to examine horizontally how these issues manifest themselves on social media platforms. However, as Trachtman observes, digital platforms present special challenges given the enhanced frequency and velocity of interaction that create novel law enforcement difficulties; and, thus, special structural and procedural rules are required to meet these challenges.

We convened the authors alongside several additional experts representing a wide cross-section of viewpoints – from Meta as well as its Oversight Board, Meedan, the National Conference on Citizenship, the *New York Times*, and the Government of Denmark – at The Fletcher School for a symposium that included a public discussion and debate of their ideas, key issues, and potential solutions. In the conclusion of this book (Chapter 15), we summarize the overarching ideas, and prescriptions for public and private policy, that have been derived from this project. With the year 2024 as the biggest election year worldwide in human history and growing fears of AI-aided disinformation, the explorations and findings covered here could not have been more timely; 2024 will, surely, alert us to the many ways in which the moderation of content – or lack thereof – on digital platforms is of enormous social and political relevance. It is our sincere hope that the chapters to follow help us gain perspectives on the depth of the challenges and pathways to possible solutions.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 01:08:42, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781009438636.001