Thomas Ehrlich, President Emeritus, Indiana University; Distinguished University Scholar, California State University, San Francisco; Visiting Professor, Stanford Graduate School of Education

Carol Greenhouse, Professor Emerita, Arthur W. Marks ’19 Professor of Anthropology, Princeton University

Davydd James Greenwood, Goldwin Smith Professor of Anthropology Emeritus, Cornell University

David Grewal, Professor of Law, UC Berkeley School of Law.

Jeffrey Hart, Emeritus Professor of Political Science, Indiana University—Bloomington

J. William Hicks, C. Ben Dutton Professor Emeritus of Law, Indiana University Maurer School of Law

Peter Katzenstein, Walter S. Carpenter, Jr. Professor of International Studies, Cornell University

T.Y. Lin, Professor of Law, National Taiwan Ocean University

Hans Lindahl, Chair of Philosophy of Law, Tilburg University, Chair of Global Law, Queen Mary University of London

Samuel Moyn, Chancellor Kent Professor of Law and History, Yale University

David J. Mullan, Emeritus Professor of Law, Queen’s University, Ontario

June Nash, Emeritus Professor of Anthropology, City University of New York

Mary Ellen O’Connell, Robert and Marion Short Professor of Law and Research Professor of International Dispute Resolution, University of Notre Dame Law School

Doris Marie Provine, Emeritus Professor, School of Social Transformation, Arizona State University

Jothie Rajah, Research Professor, American Bar Foundation

Richard Rosecrance, Professor Emeritus of Political Science, University of California—Los Angeles, Adjunct Professor of Public Policy, the Harvard Kennedy School, Senior Fellow, the Belfer Center for Science and International Affairs

Philippe Sands, Professor of the Public Understanding of Law, Project on International Courts and Tribunals, University College, London

Martin Shapiro, Professor of Law Emeritus, University of California—Berkeley, Berkeley School of Law

David H. Smith, Professor Emeritus, Indiana University—Bloomington
Ziya Umut Türem, Associate Professor, Atatürk Institute for Modern Turkish History, Boğaziçi University

Marek Wierzbowski, Professor of Law, Program Council of the Research Center on Legal Aspects of Blockchain Technology, Warsaw University

The Honorable Diane P. Wood, Circuit Judge, United States Court of Appeals, Seventh Circuit

Song Yongxin, Professor of Law, Zhejiang University, China

Elisabeth Zoller, Visiting Professor of Law, Indiana University Maurer School of Law, Professor of Public Law, University of Paris II (Panthéon-Assas)
CONTENTS

SYMPOSIUM

DIGITAL CONSTITUTION: ON THE TRANSFORMATIVE POTENTIAL OF SOCIETAL CONSTITUTIONALISM
Guest Editors - Gunther Teubner & Angelo Jr Golia

Societal Constitutionalism in the Digital World: An Introduction
Gunther Teubner & Angelo Jr Golia

Internet Bills of Rights: Generalisation and Re-Specification Towards a Digital Constitution
Edoardo Celeste

The Normative Power of Artificial Intelligence
Giovanni De Gregorio

Algorithmic Constitutionalism
Oren Perez & Nurit Wimer

Political Autonomy in the Digital World: From Data Ownership to Digital Constitutionalism
Dan Wielsch

Rage Against the Machine: Profiling and Power in the Data Economy
Irina Domurath

Digital Monetary Constitutionalism: The Democratic Potential of Monetary Pluralism and Polycentric Governance
Roxana Vatanparast

Net Neutrality: A Fundamental Right in the Digital Constitution?
Christoph B. Graber

01
25
55
81
115
131
165
197
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Against Procedural Fetishism: A Call for a New Digital Constitution</td>
<td>227</td>
</tr>
<tr>
<td>Monika Zalnieriute</td>
<td></td>
</tr>
<tr>
<td>Tackling Threats to Academic Freedom Beyond the State: The Potential</td>
<td>265</td>
</tr>
<tr>
<td>of Societal Constitutionalism in Protecting the Autonomy of Science</td>
<td></td>
</tr>
<tr>
<td>in the Digital Era</td>
<td></td>
</tr>
<tr>
<td>Raffaela Kunz</td>
<td></td>
</tr>
<tr>
<td>Robo Justice: Constitutional Issues with Judge AI</td>
<td>293</td>
</tr>
<tr>
<td>Tania Sourdin</td>
<td></td>
</tr>
<tr>
<td>The Transformative Potential of Meta’s Oversight Board: Strategic</td>
<td>325</td>
</tr>
<tr>
<td>Litigation within the Digital Constitution?</td>
<td></td>
</tr>
<tr>
<td>Angelo Jr Golia</td>
<td></td>
</tr>
<tr>
<td>STUDENT NOTES</td>
<td></td>
</tr>
<tr>
<td>The Cost of Looking Good: How Fashion and Trend-Based Consumerism</td>
<td>363</td>
</tr>
<tr>
<td>Impact The Economy, Law, and Environment</td>
<td></td>
</tr>
<tr>
<td>Allison Denton</td>
<td></td>
</tr>
<tr>
<td>The Impact of the Exhaustion of Rights Doctrine on Parallel Imports</td>
<td>391</td>
</tr>
<tr>
<td>and International Trade</td>
<td></td>
</tr>
<tr>
<td>Christian Pederson</td>
<td></td>
</tr>
<tr>
<td>Ensuring Dignity for the Survivors and the Dead: Genocide Denial</td>
<td>417</td>
</tr>
<tr>
<td>and the Law in Post-conflict Bosnia and Herzegovina</td>
<td></td>
</tr>
<tr>
<td>Daniel Schumick</td>
<td></td>
</tr>
</tbody>
</table>
Societal Constitutionalism in the Digital World: An Introduction*

GUNther Teubner** and ANGELO Jr GOLIA***

ABSTRACT

This paper introduces the symposium issue of the Indiana Journal of Global Legal Studies dedicated to “Digital Constitution: On the Transformative Potential of Societal Constitutionalism,” where a group of scholars, using societal constitutionalism as a background theory, presents concrete proposals for a digital constitutional law. The symposium issue seeks to answer three interrelated questions. What is the message of societal constitutionalism for the emerging digital constitution? How can fundamental principles of nation-state constitutions be generalized and re-specified for global digitality with a transformative outlook? What would new institutional arrangements and interpretive practices look like? In this introduction, we aim to overcome three reductive tendencies stemming from traditional constitutionalism’s legacy. We argue that digital constitutionalism needs to look beyond (1) the still dominant state-centricity of constitutional principles, (2) their exclusive focus on political power, and (3) their narrowly individualist interpretation of constitutional rights. This deconstruction opens the view to the main constitutional threats posed by digitalization—in particular, what we call the double colonization of the digital space—and to possible counterstrategies inspired by societal constitutionalism. Subsequently, we outline the content of the contributions to this symposium, grouped into four areas: (1) reformulation of constitution- and law-making; (2) digital economy; (3) institutions of constitutionalism; and (4) digital justice. Finally, we point to future developments as well as to links to other strands of literature that focus on the relationship between digital technologies and (constitutional) law.
I. EXPLORING THE TRANSFORMATIVE POTENTIAL OF DIGITAL CONSTITUTIONALISM THROUGH SOCIETAL CONSTITUTIONALISM

Traditional state-centered constitutionalism does not keep pace with the inherent dangers of the digital revolution. The digital code increases the often-discussed (self-)destructive potential of the capitalist economy, democratic politics, autonomous science, technology, and militant religion. Digitalization is accelerating the inner expansive tendencies of functional differentiation: the simultaneous politicization, monetization, scientification, and juridification of society. Surveillance capitalism, informational power, and social media’s political and religious radicalization are only some unconstrained communicative dynamics that are amplified by the application of the digital code.

Against these trends, digital constitutionalism, an emerging strand of constitutional scholarship, raises the question of whether fundamental principles of constitutionalism—notably separation of powers, democracy, fundamental rights, and rule of law—can also be established in the digital world. Such principles need to be

---

* We thank the Indiana Journal of Global Legal Studies editorial team, notably Alfred C. Aman Jr., Emma DeLaney Strenski, and Daniel Schumick. We would also like to thank the participants in the workshop “Digital Constitution: On the Transformative Potential of Societal Constitutionalism” held on 17–18 June 2022, and particularly the colleagues generously serving as discussants to the draft papers presented on that occasion: Francisco de Abreu Duarte, Lorenzo Gradoni, Amy Kapczynski, Fleur Johns, Marta Maroni, José Gustavo Prieto Muñoz, and Sofia Ranchordás. Finally, we would like to thank Anne Peters as well as the Max Planck Institute for Comparative Public Law and International Law and the research centre “Normative Orders” of the Goethe University for their institutional support. All websites last visited 8 March 2023 if not otherwise indicated.

** Emeritus Professor, Goethe University, Frankfurt.

*** Assistant Professor, University of Trento School of Law.


reformulated so that they can react to digital communication’s (self-)destructive potential. Digital constitutionalism, then, goes beyond state-centered constitutionalism in two ways. It cuts through the national/transnational as well as the state/society divide. Thus, it combines national, global, and societal perspectives.\(^5\)

Digital constitutionalism can be seen as part of the discourse on societal constitutionalism, which in the past two decades has developed within the broader galaxy of global constitutionalism.\(^6\) As a theory of legal and constitutional pluralism, societal constitutionalism has been exposed to extensive debate.\(^7\) It is indicative that one of its first comprehensive formulations used the digital sphere as a case study.\(^8\)

Up to now, societal constitutionalism has been mainly used as an analytical framework to frame issues of digital constitutionalism.\(^9\) However, its normative and transformative dimensions still need to be explored. Indeed, the relevant literature has rarely developed concrete

---


6. See GUNther Teubner, CONSTITUTIONAL FRAGMENTS: SOCIETAL CONSTITUTIONALISM AND GLOBALIZATION (Oxford University Press, 2012) (detailing two monographic treatments); see also POUL F. KJAER, CONSTITUTIONALISM IN THE GLOBAL REALM: A SOCIOLOGICAL APPROACH (Routledge, 2014); see also Angelo Jr Golia & Gunther Teubner, Societal Constitutionalism: Background, Theory, Debates, 15 ICL - VIENNA J. INT'L. CONST. L. 357 (2021) (taking into account criticisms and debates); see also Antje Wiener et al., Global constitutionalism: Human rights, democracy and the rule of law, 1 GLOB. CONSTITUTIONALISM 1, 7 (2012) (explaining global constitutionalism); see also NEIL WALKER, INTIMATIONS OF GLOBAL LAW (Cambridge University Press, 2014).


9. See, e.g., EDOARDO CELESTE, DIGITAL CONSTITUTIONALISM: THE ROLE OF INTERNET BILLS OF RIGHTS (Routledge, 2022); see also Lorenzo Gradoni, Constitutional Review via Facebook’s Oversight Board. How platform governance had its Marbury v Madison, VERFASSNUNGSBLOG ON MATTERS CONST. (Feb. 10, 2021), https://verfassungsblog.de/foh-marbury-v-madison; see also Nofar Sheffi, We Accept: The Constitution of Airbnb, 11 TRANSNAT'L LEGAL THEORY 484 (2020); see also Gill et al., supra note 4; see also Marco Bassini, Fundamental Rights and Private Enforcement in the Digital Age, 25 EUR. L. J. 182 (2018).
legal policies based on that framework. The risk is to render both societal and digital constitutionalism incapable of a critique that engages with societal power in the digital sphere and sets clear normative standards.

The primary goal of this symposium is to explore the transformative potential of digital constitutionalism through the lens of societal constitutionalism. Using societal constitutionalism as a background theory, a group of scholars presents concrete proposals for a digital constitutional law. In this way, this symposium issue seeks to answer three interrelated questions. What is the message of societal constitutionalism for the emerging digital constitution? How can fundamental principles of nation-state constitutions be generalized and re-specified for global digitality with a transformative outlook? What would new institutional arrangements and interpretive practices look like?

In this introduction, we aim to overcome three reductive tendencies stemming from traditional constitutionalism’s legacy. We argue that digital constitutionalism needs to look beyond (1) the still dominant state-centricity of constitutional principles, (2) their exclusive focus on political power, and (3) their narrowly individualist interpretation of constitutional rights. This deconstruction opens the view to the main constitutional threats posed by digitalization—in particular, what we call the double colonization of the digital space—and to possible counterstrategies inspired by societal constitutionalism. Subsequently, we outline the content of the contributions to this symposium, grouped into four areas: (1) re-formulation of constitution- and law-making; (2) digital economy; (3) institutions of constitutionalism; and (4) digital justice. Finally, we point to future developments as well as to the links to other strands of literature that focus on the relationship between digital technologies and (constitutional) law.

II. DIGITAL CONSTITUTIONS BEYOND TRADITIONAL CONSTITUTIONAL THEORY

Against state centricity. To be sure, traditional state constitutionalism still has considerable potential to protect against digital authoritarianism in the political system. China’s “Social Credit System”¹⁰ as well as US predictive policing,¹¹ both of which use digital

---

technologies to suppress potential threats to state power, are exemplary cases. In order to preserve the democratic potential of digital technologies against repressive politics, state constitutions need to establish new protective rules (e.g., free and continuous access to the internet and preservation of anonymity under certain conditions).\textsuperscript{12} However, it is wrong to reduce digital constitutionalism to a set of rights, governance norms, and limitations on the exercise of the states’ power on the internet. State-centric constitutionalism fails to address the collective power exercised by private actors. To stand against repressive tendencies in non-state sectors of society, (i.e., in market transactions) formal organizations, or transnational regimes constitutional protection needs to reach far beyond the power threats of the state world.

Today, the digital space is the new non-state sector of global society that needs comprehensive constitutionalization. This does not only require new constitutive rules— that is, complex institutional structures sustaining the emergence and action of relevant actors, including higher-level normative texts\textsuperscript{13} and intricate networks of organizations and procedures.\textsuperscript{14} Even more urgent is the need for new limitative rules, produced by both state and non-state law, directed against the digital power of private actors, notably anticompetitive practices of Silicon Valley and Big Tech companies.\textsuperscript{15}

---


\textsuperscript{13} See CELESTE, supra note 9.

\textsuperscript{14} Cf. Oren Perez & Nurit Wimer, \textit{Algorithmic Constitutionalism}, 30 IND. J. GLOB. LEGAL STUD. (2023);and Nofar Sheffi, \textit{We Accept: Bit-by-Bit Constitution} (forthcoming).

\textsuperscript{15} See Samuel Stolton, \textit{EU braces for Big Tech’s legal backlash against new digital rulebook}, POLITICO (Aug. 10, 2022, 6:00 AM), https://www.politico.eu/article/eu-brace-
Beyond (social) power. However, there lurks a second and more subtle reductionism. While digital intermediaries, as the new centers of power beyond the state, are the legitimate target of political critique, it is not sufficient to focus exclusively on power in this context.\(^1\) The preoccupation with social power obscures the excesses of other expansive communication media (money, knowledge, and juridical authority)\(^2\) that—even in situations when social power centers are absent—require constitutional limitation.\(^3\) Constitutional strategies ought to develop limitative rules not only against the negative externalities produced by the power imperative of politics but also—and particularly—against the externalities of the profit imperative of the economy, the reputation imperative of science, the innovation imperative of technology, the news imperative of information media, the health imperative of the medical system, and the juridification imperative of the law.\(^4\)

Digitality itself is the new communicative medium against whose externalities constitutional protection is needed. Most conspicuous among them is the trend that digital technology creates its own social realities. This “hyperreality” has the potential to monopolize

---


\(^2\) Here understood not as coercion or merely as self-interested influence on social actors’ behavior but rather as a specific communication medium (see infra note 19 below) that makes accepting Alter’s actions as the premises of Ego’s actions probable. In the functionally differentiated society, power is the specific medium of the political system. It can potentially also be realized in other systems, but without being able to obtain the capacity to reproduce that it has in politics. Indeed, power reproduces itself in the form of obedience to a command. In other words, it is realized when the action sequence command-obedience is combined with a sequence of threat of sanction (if you do not obey, I will punish you); cf. CLAUDIO BARALDI ET AL., *UNLOCKING LUHMANN. A KEYWORD INTRODUCTION TO SYSTEMS THEORY* 175 (Bielefeld University Press. 2021).

\(^3\) Rudolf Stichweh, *Systems Theory*, in *INTERNATIONAL ENCYCLOPEDIA OF POLITICAL SCIENCE* 1, 6 (Bertrand Badie et al. eds., 2011) (Understood as the “effect mechanisms” of the functionally differentiated society. Communication media “. . . are based on symbols which are thought to be effective in communication – e.g. symbols of money, power, truth or love -, and which as such effective symbols motivate other social actors to do something they would not have done without this effective use of symbols.”)


communication in other (life-)worlds, and totalize its own reality construction at the expense of other ones.20 “With the virtual, we enter not only upon the era of the liquidation of the real and the referential, but that of the extermination of the other.”21 In its relation to law, the digital code creates autonomous normative orders and tends to undermine the law’s normative order. Since the rigid calculations of algorithms induce a fusion of unilateral rule-making, rule-application, and rule-enforcement, they risk destroying the civilizing and humanizing aspects of the rule of law, particularly the hermeneutics of legal argumentation.22

Beyond individual rights. The third reductionism—the exclusively individualist dimension of constitutional rights—of traditional constitutionalism comes in. Of course, a Bill of Rights for individual users of social networks is important in combating digitality’s damaging effects on privacy, mental health, and political engagement of citizens.23 “Digital vulnerabilities” is a political project exploring how digital technologies exacerbate pre-existing human vulnerabilities or create new ones.24 However, “the real fundamental rights issue lies at the trans-individual, discursive level: Platforms are expansive social systems, which may thwart society's autonomous self-reproduction,” (emphasis added).25 Thus, societal constitutionalism, reaching beyond the individual dimension, focuses on the equally crucial institutional


dimension of constitutional rights. This means protecting the integrity of vulnerable social configurations and less powerful collective actors (protest movements, trade unions, independent media, and education and research institutions) against their encroachment. For instance, “as a fundamental right, academic freedom guards the individual autonomy of the scholar, but also facilitates functional differentiation of societal systems, in this case, protecting science against unwanted intrusions from politics, economics or religion.”

And similarly vulnerable institutions are emerging in socio-digital spaces as well (e.g., Wikipedia, the open-source movement, digital commons, public software repositories, and social movements of platform workers) all of which, in their still fragile autonomy, require constitutional protection.

III. DOUBLE COLONIZATION, THE NEW DIGITAL POLITICAL ECONOMY, AND COUNTER-STRATEGIES: RESISTIBILITY AND CONTESTABILITY

Up to this point, we have focused on constitutional problems created by digital technology itself. In addition, societal constitutionalism identifies future threats that lurk elsewhere—that is, in the negative effects of digitalization on the polycontextural structure of contemporary society. When the two dominant functional systems, politics and the economy, are digitalized comprehensively, their surplus pressures, profit and power, are massively reinforced by the equally strong surplus pressures of digitality. Internally, digitalization intensifies the endogenous growth dynamics in the political and economic systems.Externally, it aggravates their expansive tendencies directed toward other social systems. Both trends result in the double colonization of the digital space: the power-profit complex produces a digital


28. See generally Gotthard Günther, Life as Poly-Contexturality, in BEITRÄGE ZUR GRUNDLEGUNG EINER OPERATIONSFAHIGEN DIALEKTIK 283–306 (Gotthard Günther ed., 1976) (We understand “polycontexturality” as a characterizing feature of modern societies: the plurality of mutually irreducible social perspectives. They are incompatible with one another and can be overcome only by rejecting certain values, which in turn leads to different binary distinctions).

29. See generally LESSIG, supra note 22 (discussing early warnings against the dangers stemming from the dominant coupling of ‘government’ and ‘commerce’).

30. Cf. WANG, supra note 20, at 6 (describing the Habermasian perspective).
totalitarianism which impedes the potential plural evolution of
digitality. The complex reduces the polycontextural structure of the
digital space to a duopoly: a “public sector” driven by digitalized power
and a “private sector” driven by digitalized profit. This dominating
duopoly—which could be called the new “digital political economy”31—
has the potential of structurally corrupting the new but still fragile
socio-digital institutions32 which are emerging in the other social
domains: science, education, health, and art. The new dystopia is the
fusion of homo oeconomicus and homo politicus into homo digitalis.33 In
normative terms, combating this dystopic scenario employs three
constitutional strategies.

The first strategy is to develop constitutional restraints on
digitalized politics. As said above, when digitality reinforces state
power, the repressive potential of political systems is intensified.
Against that, traditional state constitutionalism will have to indeed
fulfill its century-old promise: consolidate the rule of law, strengthen
constitutional rights, and fight anti-democratic practices. However, for
the new dangers of digitalized state power, key principles of state
constitutionalism need to be transformed into their societal counterpart.
Their generalization into broader constitutional strategies and their re-
specification for digital power is needed. Digital constitutionalism will
have to put severe limits on biometric (especially “crowdwashed”)
surveillance,34 automated decision-making,35 and “hyperfudge”
technologies which use Big Data for regulation by design.36
The second strategy is to set constitutional limits on algorithmic profit-making. The economization of the digital medium is the blind spot of traditional constitutionalism which focuses on state power. Digital technologies have increased the commodification tendencies of global capitalism. The combination of oligopolistic markets, enhanced computational technologies and data-driven business models has dramatically expanded the possibility for economic actors to affect society through the code of economic accumulation. In particular, the strategies of informational capitalism combine the pressures of the profit motive with digitality’s pressures of attention maximization. Empirical studies have produced “powerful observational evidence of destructive dynamics, including the fast diffusion of misinformation, manipulation campaigns, ideological (self-)segregation, and extremism” which are produced by digital attention maximization. This requires new constitutional measures against the technological network protocols that are self-enforcing.

Today, constitutional counter-strategies are slowly emerging. The Digital Services Act (DSA) adopted by the European Union is a significant regulatory experiment whose concrete effects will only emerge in the future. It prohibits the use of UX tweaks to manipulate or force consent and requires platforms to offer parity in consent flows for refusing or agreeing to hand over data (Art. 25); ad profiling of minors (Art. 28); and the use sensitive personal data (such as racial or ethnic origin, political or religious affiliation, and sexuality or health data) for behavioral targeting (Art. 26 para. 3). An additional important element in the DSA is data access and scrutiny into algorithmic background operations, in order to make transparent how the profit motive overlays “innocent” technical operations. In this sense, the DSA moves towards constitutional restraints of the digitalized economy when it requires “very large online platforms” to periodically conduct and publish assessments concerning systemic risks, particularly before launching new services (Art. 34), with related mitigation obligations (Art. 35), regulatory oversight of their algorithms, and to provide public interest researchers with access to data to enable independent scrutiny of platform effects (Art. 40).

Besides the features of specific instruments such as the DSA, dispute-settlement bodies need to start reviewing the “private”

SARTOR, ALGORITHMS AND REGULATION 131 (Micklitz et al. 2021); [Author’s name] supra note 11, at 131–38.
37. See generally COHEN, supra note 2.
38. González-Bailón & Lelkes, supra note 3, at 165.
regulations of the “digital governors”\textsuperscript{40} with far-reaching constitutional controls, mirroring the constitutional review of state legislation performed by state courts.\textsuperscript{41} In this sense, it is crucial to frame the digital platforms’ terms of use not simply as standardized contracts but as forms of unilateral law-making power.\textsuperscript{42} Indeed, platforms increasingly exercise their power “through non-negotiable, one-sided and deliberately opaque ‘terms of service’ contracts.”\textsuperscript{43} Therefore, courts need to impose strict scrutiny standards on digital private government regimes. Most importantly, they must rely on more than legal consent of individuals because it does not take into consideration problems of asymmetric information, unequal bargaining power, and collective negative externalities.\textsuperscript{44} Thus, the private law control over the fairness of contractual provisions is—has to be—transformed into a constitutional review of the non-state law emerging in the digitalized economy, a new and increasingly important type of constitutional review performed by both state and non-state judicial bodies. From this perspective, strategic litigation, activated by both individuals and collective actors, based on both state and non-state rules, is an additional instrument to trigger the emergence of both constitutive and limitative constitutional norms within the digital sphere.\textsuperscript{45}

Moreover, antitrust law needs to come in and develop constitutional rules to protect the integrity of information processes within digital networks (e.g., prohibiting unfair “dark patterns” and other


\textsuperscript{42} For an exemplary case going in this direction, see Tribunal of Rome, 5 Dicembre 2022, n. 17909/2022 (It.).


\textsuperscript{44} See Patrik Hummel et al., Own Data? Ethical Reflections on Data Ownership, 34 Phil. & Tech. 545 (2021); see generally Ignacio Cofone, Beyond Data Ownership, 43 Cardozo L. Rev. 501 (2021); Martin Tisné & Mariette Schaake, The Data Delusion: Protecting Individual Data Isn’t Enough When the Harm is Collective, Luminare (2020).

\textsuperscript{45} Cf., Vera Strobel, Strategic Litigation and International Internet Law, in Digital Transformations in Public International Law 261-284 (Angelo Jr Golia et al. eds., 2022) (in the field of public international law); see generally Angelo Jr Golia, The Transformative Potential of Meta’s Oversight Board: Strategic Litigation within the Digital Constitution? (for the non-state normative system of Meta) (forthcoming).
manipulative digital practices). Another proposal, inspired by northern European corporate law models, would establish forms of co-decision with external representatives of collective interests (health, science, and environment) within the corporate boards of digital service providers. All of this would amount to a societal constitutionalism “by procedure” imposing a mandatory procedural framework for the self-limitation of digital processes. Here again, another instrument of the European Union’s recent regulatory efforts, the Digital Markets Act (DMA), presents interesting developments, notably the provisions involving third parties and competitors in the monitoring and compliance functions (Arts. 27 and 28).

The third strategy focuses on the institutions of a “digital civil society.” It aims to protect the non-state/non-profit sectors of the digital space in constitutional terms. To guard against the double colonization by profit and power, the integrity of science, journalism, education, medicine, and art needs constitutional protection. Thus, the digital third sector requires constitutive rules for developing stable socio-digital institutions: hacker communities, digital NGOs, digital commons, Wikipedia, and open-source. However, the so-called tragedy of the digital commons reveals self-destructive tendencies even within digital civil society. The average user of information technology is exploiting common resources until they can no longer recover. Users do not pay attention to the consequences of their behavior. The tragedy of the digital commons has a second consequence—the pollution of the infosphere (i.e., the indiscriminate and improper usage of technology and digital resources and the overproduction of data). Excess information leads to corruption of communication and information overload. Both tendencies are within the legitimate field of

46. See generally Corina Cara, Dark Patterns In The Media: A Systematic Review, 7 NETWORK INTEL. STUD. 105 (2019). As an interesting recent development, see Judgment of the Court of Justice of the European Union in Case C-252/21, Meta Platforms and Others (General terms of use of a social network), 4 July 2023 holding that a national competition authority can find, in the context of the examination of an abuse of a dominant position, that the data protection rules have been infringed.


49. See generally Raffaela Kunz, Threats to Academic Freedom under the Guise of Open Access (Mar. 8 2022), https://verfassungsblog.de/threats-to-academic-freedom-under-the-guise-of-open-access/ (Protecting science, especially from the perspective of international human rights law); see generally Verschraegen, supra note 26, at 164.

constitutional self-limitation which needs to be supported by external pressures from politics and civil society.

But most importantly, the digital third sector needs strong rules against the external pressures of both profit and power surpluses. The integrity of science is affected by negative externalities produced by the digital political economy when digitalization and the so-called economic reinterpretation of open access are increasing publish-or-perish reputation-seeking dynamics, predatory publishing, and reinforcing the position of hegemonic actors in science. Large publishing companies, like Elsevier, Wiley, and Springer are in a position to invisibly and strategically “exert control, over key university decisions—ranging from student assessment to research integrity to financial planning.”

Likewise, the integrity of journalism is endangered by real-time web analytics, clickbait, and information bubbles, dynamics that have already led to important changes: consolidation of larger news organizations and transformation of the professional self-understanding and self-organization of journalism.

The constitutional counterstrategies inspired by societal constitutionalism are expressed by two key concepts: resistibility and contestability. They represent two sides of a coherent strategy against the double colonization of the digital space by the power-profit complex. This strategy has the potential to transform digital constitutionalism from an academic concept into a socio-political movement. Resistibility implies civil society's defense against the digital political economy. Against the colonizing tendencies of digitalized politics, it will have to create or sustain social counterpowers, mainly by protest movements and civil society groups. This is not just wishful thinking. Indeed, “the use of algorithmic governance in increasingly high-stakes settings has

51. See Raffaela Kunz, Opening Access, Closing the Knowledge Gap?, 81 HEIDELBERG J. INT'L L., 23, 43–45 (2021); see generally Debate Open/Closed, VERFASSUNGSBLOG ON MATTERS CONSTITUTIONAL, https://verfassungsblog.de/category/debates/open-closed/.

52. CLAUDIO ASPESI ET AL., SPARC, LANDSCAPE ANALYSIS 5 (2019).


55. Celeste, supra note 23 (forthcoming 2023) (observing (correctly) that numerous proposals for digital Bill of Rights, which have emerged in recent years, are indicators for a social movement that produces constitutional counterstrategies).
generated an outpouring of activism, advocacy, and resistance."  

Against the excessive economization of the digital world, profit-threatening strategies are the most promising instruments which law and politics could impose. Contestability will imply, internally, the protection of self-contestation. Digital platforms will have to allow procedures for internal opposition and whistleblowing. Externally, the expansion of access to justice is needed, against algorithmic politics and digitalized economization. Ultimately, this symposium issue appeals to "institutional imagination" in the sense of Roberto Unger.  

It has a critical, normative, and transformative outlook and aims to offer concrete proposals within the broader context of digital constitutionalism.

IV. THE SYMPOSIUM’S CONTRIBUTIONS ON FOUR MACRO-ISSUES: CONSTITUTION-MAKING, DIGITAL ECONOMY, INSTITUTIONS OF CONSTITUTIONALISM, AND DIGITAL JUSTICE

We now briefly outline the content of the individual contributions of this symposium issue. All the authors had already engaged with issues of digital law and politics. Expanding on their previous works, they address crucial issues of digital constitutionalism through the lens of societal constitutionalism and come up with concrete proposals. In particular, the contributors examine experimental approaches. Via case studies in different fields, they point to shortcomings and work out alternatives. Moreover, they critically reflect on how new solutions impact both hegemonic and subaltern positions affected by digital technologies. The symposium issue is organized into four sections, each addressing substantive and procedural problems of digital constitutionalism.

The first section addresses constitution-making by the digital code. Edoardo Celeste analyzes the potential of so-called “Internet Bills of Rights.” The Bills of Rights generalize and re-specify constitutional norms in the digital sphere, creating the transformative potential of societal constitutionalism. In particular, Celeste highlights how, even though they are not legally binding sources, they represent a flexible instrument whereby their promoters are free to experiment with new legal solutions gradually and more democratically, including actors...
beyond the worlds of politics and business.

Giovanni De Gregorio focuses on the digital code as a matrix of constitutional normativity and deals with it in the general framework of constitutionalism as a normative project. Starting from the observation that artificial intelligence (AI) systems create their own norms by defining generative layers of normativity in the algorithmic society, he argues that automated decision-making systems autonomously develop norms by experience and learning within an opaque technological space that tends to escape the logic of the rule of law. Within this context, he discusses the European Union’s proposed Artificial Intelligence Act as an example of how the rule of law can limit delegation in the digital age.

Oren Perez and Nurit Wimer also address the constitutional impact of AI on regulation, but they focus on the content moderation of digital platforms. They examine the Facebook content moderation regime, already partially controlled by algorithms. Starting from a critique of current approaches based on ethical engineering, they develop “algorithmic constitutionalism” as an original approach to AI governance. They demonstrate how it can be applied to the Facebook content moderation regime and describe the difference between societal and algorithmic constitutionalism. Indeed—and paradoxically—the attempt to subject the AI algorithm to external control opens the door for the AI agent to intervene in that process, potentially undermining its very purpose. Finally, they explore the implications of their argument for the DSA.

The second section deals with the politics of data property and the law’s role in shaping the interface between economy and digitality. Dan Wielsch observes that, in contemporary economic systems, data are taking their place alongside labor and capital, which raises questions about the need and the legitimacy of creating exclusive rights to data or “data ownership.” However, legal theory does not only have to develop an adequate concept of “data” and to explicate the social functions of related property rights. It also has to align a potential data ownership with the broader idea of social ordering through property rights, taking into consideration the normativity of social orders constituted through the exercise of rights and ensuring that those affected by such orders can participate in shaping them. Concerning the function of individual

59. De Gregorio, supra note 22.
rights for social practice—he argues—two further questions arise: the implications of the normativity of this practice for the rights and, correspondingly, the participation of the right holders in social practice. Ultimately, and to the extent private rights would allow for changing the rules of social order, they become political rights.

Irina Domurath deals with algorithmic profiling as an example of datafication and machine colonisation. She examines the emergence of an EU digital constitution from the perspective of societal constitutionalism. Via an internal critique of societal constitutionalism, she questions its assumptions concerning the capacity of societal actors and non-legal media, such as public outrage and litigation, to exert the pressure needed for changes from within. She resorts to the insights of the emerging “Law and Political Economy” (LPE) scholarship to understand the structural power of companies that inhibit the build-up of external pressure and to justify the adoption of a counter-concept of structural digital vulnerability.

LPE scholarship is a point of reference also for Roxana Vatanparast, who addresses yet another side of the relationship between digital technologies and the economy: digital money. Focusing on its governance and democratic potential, she explores the opportunities provided by the pluralism of digital money and polycentric governance to embed values that might otherwise not be valued in market societies. In particular, she refers to two case studies—namely digital money built by and for stateless populations utilizing blockchain technology and digital fiat currency that has the privacy-preserving features of cash and promotes financial inclusion. She argues that the pluralism of digital currencies that utilize public and not-for-profit institutional architectures has a higher democratic potential than forms of digital money driven by profit and extractive motives.

The contributions of the third section explore different ways to reframe the fundamental institutions of constitutionalism (rights, democracy, separation of powers, procedures) in the digital sphere. Using net neutrality as a case study, Christoph Graber argues for a reconstruction of fundamental rights as institutions. They should bundle normative expectations related not only to the protection of individual positions but also to the defense of institutional autonomies.

63. See Pistor, supra note 31, 183–204; see Kapczynski, supra note 2.
against society’s self-destructive tendencies. Starting from existing statutory guarantees of net neutrality in certain jurisdictions, he argues for the development of constitutional structures and processes—the next step to be expected according to the theory of societal constitutionalism. From a normative perspective, he explores how net neutrality protection should be institutionalized as a fundamental right. In particular, he argues that two preliminary issues need to be addressed: first, how to adequately conceptualize the relationship between the social and the technological, and second, how fundamental rights should be conceived beyond state-centrism. He concludes that a sociological reflection of fundamental rights as institutions of society will serve as a benchmark for evaluating future developments of constitutional legal doctrine.

Monika Zalnieriute takes a different path. Resorting to critical and decolonial scholarship, she questions proceduralist solutions often offered in the digital constitutionalism scholarship. She criticizes what she calls “procedural fetishism” as a strategy of digital imperialism to hide and reinforce US dominance, colonial exploitation, and environmental degradation. A new digital constitution—she argues—would have to shift its focus from procedural and soft law initiatives toward substantive accountability and tangible legal obligations of the tech companies. Even more urgently, digital constitutionalism needs to recognize colonial practices of extraction and exploitation, paying attention to the voices of indigenous communities of the “Global South.” Only with these mutually reinforcing efforts will a new digital constitution debunk corporate and state agendas of procedural fetishism and establish a new social contract for the digital age.

Raffaela Kunz addresses such issues from yet another perspective. Focusing on Open Science as a case study, she observes how awareness about the dark sides of digital technologies has been rising in recent years. Leading academic publishing companies have long started to tap into the data analytics business, with negative consequences for the consolidation of an oligopoly in the academic publishing industry and large-scale corporate influence on science. Against this background, she argues that traditional constitutionalism cannot capture the subtle yet systemic risks that science faces in the digital age. Societal constitutionalism is not only a useful analytical lens but also helps to

---

respond to these threats. It provides valuable lessons for debates on digital constitutionalism and the effective protection of fundamental rights in the digital age.

The fourth and final section examines the interface between digital technologies and judicial adjudication, both in the private and public spheres. Tania Sourdin explores AI’s problematic impact on states’ judicial functions. She observes how the emerging relations between judges, courts, and AI technologies challenge conventional governance theory, as they require focus on social interaction to explore how judicial responsiveness might support the development of ethical approaches that care for vulnerable populations. Societal constitutionalism—she argues—opens a new approach to justice, which promotes human well-being, which in turn supports disruptive technologies in the justice sector. She also reflects on the challenges presented by this approach, which are readily apparent in justice conceptions that focus on “fast,” “low cost” justice delivery, in the absence of justice itself.

In the final contribution, Angelo Jr Golia focuses on the Oversight Board (OB), the independent adjudicative body established by Meta to make consequential precedent-setting content moderation decisions on Facebook and Instagram. He proposes a potential strategy to make digital platforms responsive to external demands concerning their broader societal impact. Starting with an analysis of Meta’s normative system from the perspective of societal constitutionalism, he assesses the actual extent of juridification and constitutionalization. Aiming to put informational/surveillance capitalism “on trial,” he then outlines a litigation strategy concerning the mental health distress of young people and, by these means, ultimately uses Meta’s normative system to thematize broader, systemic effects of social networks in constitutional terms.

V. Red Threads: Definitional Issues, Materiality and Conflict, Regulatory Approaches, Systems Theory

One of the most important and challenging tasks of digital constitutionalism is to make visible convergence for different strands of scholarship that deal with the impact of digital technologies on fundamental rights, democracy, and rule of law. Such a convergence is possible especially with those approaches which do not speak—at least, not explicitly—the language of constitutionalism. This opens broader conversations, both within and outside of constitutional scholarship.

68. See Sourdin, supra note 22.
69. See Golia, supra note 45.
Against this background, in this final section, we highlight some red threads and link them to parallel debates. In particular, four red threads are emerging, concerning definitional issues, digital materiality and conflict, regulatory approaches, and systems theory.

A first red thread relates to digital constitutionalism’s identity. Indeed, its inner complexity is well reflected in its own (lack of shared) definition.⁷⁰ The contributions to this symposium reveal a variety of meanings and normative orientations in the definition of “digital constitution” and “digital constitutionalism.”⁷¹ In contrast to over-hasty critique,⁷² these nuances—this internal ambiguity, if you want—is not necessarily negative nor is it an instrument to co-opt the symbolic capital of constitutionalism. Instead, it allows multiple discourses with a shared normative outlook to co-exist, interact, and potentially compensate each other’s limits. This internal variety also facilitates critical engagements and contributes to debunking co-option attempts.⁷³ What matters is the shared ambition to build legal instruments that protect and constrain the dynamics of the digital code in its relation to power, money, faith, and juridical authority. To formulate it in a more traditional way, the goal is the translation and implementation of constitutional principles into different, newly emerged societal fields. However, the normative horizon of any constitutionalism, be it state-centered, global, or societal, defines the project.⁷⁴

The project can be seen as both, a conceptually ambitious “digital constitution” or a more modest “constitutionalism for the digital.” A serious engagement with digital technologies presupposes consideration of a plurality of normativities, which variably interact, overlap, conflict, and influence each other. Only within this process may a “digital constitution” emerge. But—importantly—such emergence is contingent and by no means necessitated. It needs to be actively pursued by

---

70. See Celeste, supra note 4 (for an overview of the definitional issues); and more recently, De Abreu Duarte at al., supra note 5.

71. See De Gregorio, supra note 22; Zalnieriute, supra note 66; Perez & Wimer, supra note 14; Domurath, supra note 24.


74. See De Abreu Duarte at al., supra note 5.
different strategic actors, including engaged scholars. Any concept of transformative digital constitutionalism will aim at laying the analytical and normative preconditions for such emergence.

There is a second red thread: (digital) materiality and conflict. Indeed, thinking digitality through the lens of societal constitutionalism allows for linkages with the material preconditions of the digital constitution in two senses. First, the socio-technical substratum of digital technologies influences its constraining effects, the actual possibilities for transformation, and the contestation of norms and policy solutions. Thus, technology facilitates as well as hinders the emergence of constitutional norms. Second, societal constitutionalism requires looking at the concrete—indeed, material—socio-economic relationships sustained by the legal infrastructure of the digital ecosystem. The processes of value extraction amplified by digitality, the redistributive effects at national and global levels, and the capacity of politics, science, and law as distinct social fields to resist colonization by economic rationality are all points of contact with LPE scholarship, as several contributions of this symposium issue show. As mentioned above, countering the negative effects of the conflation between a “public sector” driven by digitalized power and a “private sector” driven by digitalized profit—what we called the new “digital political economy”—should be one of the transformative goals of a digital law informed by societal constitutionalism.

Here, we submit that the specific contribution of societal constitutionalism lies in polycontexturality as one of its analytical starting points. Indeed, against the risk of exclusively focusing on the economy/politics interface, societal constitutionalism insists on the multiplicity of mutually irreducible social perspectives reproduced by digitality and their collisions. In normative terms, this calls for fine-tuned approaches that account for the specific dynamics of different social systems—among them, law, science, and religion—thus orienting normative and policy solutions towards the reflexivity specific to each social field.

75. Graber, supra note 65, at 8; see also Christoph B. Graber, How the Law Learns in the Digital Society, 3 L. TECH. & HUM. 12, 22 (2021).

76. Wielsch, supra note 61, at 3; Domurath, supra note 24, at 2; Kunz, supra note 67, at 8; see Vatanparast, supra note 64, at 1; see Ioannis Kampourakis, Bound by the Economic Constitution: Notes for “Law and Political Economy” in Europe, 1 J. L. & POL. ECON. 301 (2021) (for an excellent contribution linking LPE and societal constitutionalism).

77. See Günther, supra note 28.

78. See especially Kunz, supra note 67, at 20; Graber, supra note 65; IMMACULATE D. MOTSI-OMOLJADE, CRYPTOCURRENCY REGULATION. A REFLEXIVE LAW APPROACH (Routledge, 1st ed. 2022).
This leads to a third red thread, namely contribution to the regulation scholarship. This symposium helps dispel some mischaracterizations of societal constitutionalism. It is plainly wrong to say that societal constitutionalism only focuses on private ordering for the emergence of constitutional normativity. It is similarly wrong to assert that it is inspired by a neo-liberal ideology that legitimizes private powers and supports exclusively private self-regulation and the retreat of states’ regulation. In contrast, societal constitutionalism calls for the inclusion of normativities emerging from all social fields, including state-based politics. Importantly, this does not necessarily mean “less government.” Rather, even state regulation, if oriented towards effective constitutionalization, needs to be translated into self-constitutionalization. In order to perform its constituting and limiting functions, digital constitutions need to respond to the specific communicative structures/processes of digitality. More concretely, legal-political rules need to be reconstructed by the digital code. In other words, regulatory strategies aimed at an effective constitutionalization of the digital sphere may require more or less state regulation. However, such constitutionalization cannot be based exclusively on politically legitimated norms, even when they derive from authentically deliberative processes. In the end, societal constitutionalism asks for a strategic interaction of qualitatively different kinds of norms, as affected by the digital code. From this perspective, it is not surprising that several contributors—adopting insights coming from societal constitutionalism—called for a more significant role for states, ranging from the expansion of their positive obligations to the establishment of clearer and “harder” prohibitions. Likewise, the problematization of the rule of law—a classic principle of modern constitutionalism—formulated by several contributions and the strategies for its re-

81. Cf. Golia & Teubner, supra note 6, at 388.
82. Id. at 388–95.
83. Cf. Celeste, supra note 23, at 4; De Gregorio, supra note 22, at 3; Perez & Wimer, supra note 14, at 1 Golia, supra note 45, at 8.
85. Cf. De Gregorio, supra note 22, at 1; Zalnieriute, supra note 66, at 23.
86. Cf. De Gregorio, supra note 22, at 1; Zalnieriute, supra note 66, at 1; Perez & Wimer, supra note 14, at 1; Sourdin, supra note 22, at 1.
specification in different contexts are another example of how societal constitutionalism contributes to combining distinct normativities, principles of legitimacy, and regulatory approaches. As mentioned above, the normative effects stemming from digital technologies and algorithms need to be reconciled with the rule of law in a different way from what happened in “analog” constitutionalism.\textsuperscript{87} In positive terms, the contributions show the importance of looking for solutions linking the coercive effects of technology\textsuperscript{88} with the normative structures and processes that are specific to law\textsuperscript{89} and its human features.\textsuperscript{90}

The fourth and final red thread emerging from the symposium is the relationship with systems theory. Does systems theory provide an analytical framework for a digital constitution?\textsuperscript{91} Societal constitutionalism—as developed in the last two decades—builds on Luhmann’s theory of social systems and, at the same time, thinks in the normative terms of constitutionalism. Some contributions of this symposium show that Luhmann’s theory of functional differentiation opens new perspectives for a transformative reconstruction of digitality.\textsuperscript{92} The contributions reveal a further aspect of systems theory which thematizes the risks of datafication.\textsuperscript{93} Here, it is an open question of whether constitutional theory should concentrate only on the effects that the datafication has on already-existing communication media\textsuperscript{94} or on the effect it has on the digital code itself, as a new communication medium. This is the perspective of “bit-by-bit-constitutionalism” which identifies constitutional processes in the digital architecture itself.\textsuperscript{95} From a more speculative perspective, the symposium calls for socio-legal...

\textsuperscript{87} See esp. Perez & Wimer, supra note 14.
\textsuperscript{88} See generally COHEN, supra note 2.
\textsuperscript{92} See Kunz, supra note 67, at 16; Graber, supra note 65, at 5.
\textsuperscript{93} See Domurath, supra note 24, at 1; Kunz, supra note 67, at 19; Graber, supra note 65, at 9.
\textsuperscript{94} See Stichweh, supra note 17, at 6.
\textsuperscript{95} Sheffi, supra note 14.
research to investigate whether the impact of digitalization is so significant that it will trigger a departure from functional differentiation as the primary form of societal organization.96 Put otherwise, data-driven social processes such as the resocialization of power—the capacity of non-political, non-state collective actors to increase the probability of Alter’s acceptance as the premises of Ego’s actions—may go as far as to provoke the emergence of new, unprecedented forms of societal differentiation.97 These questions certainly cannot be answered only by constitutional theory. However, such a broadening of the horizon is necessary for any constitutionalism that aims to rise to the level of complexity required by the involved issues and, potentially, to offer normative solutions for a digital constitution. Perhaps, the symposium will contribute to this debate.

96. See BARALDI ET AL., supra note 16, at 65–70 (which replaced segmentation and stratification in modern societies).
Internet Bills of Rights: Generalisation and Re-Specification Towards a Digital Constitution

EDOARDO CELESTE*

ABSTRACT

Scholars have advocated the need for a “digital constitution,” including the idea of drafting an “internet bill of rights.” Numerous civil society groups have crafted decalogues of digital rights, giving rise to a movement advocating a form of digital constitutionalism. This paper investigates the added value of these texts is in the constitutional ecosystem. These declarations highlight the transformative potential of societal constitutionalism well. By adopting the language of constitutions, they seek to be part of the current conversation on how to translate the core values of contemporary constitutionalism in the context of the digital society through a process of generalisation and re-specification of principles. Internet bills of rights are not legally binding sources, yet represent a ductile instrument whereby their promoters are free to experiment with new legal solutions in a gradual way and in a more democratic manner. This includes actors beyond the worlds of politics and business.

I. INTRODUCTION

We need a “digital constitution.” This is a claim that has frequently resonated in the words of various digital literati, academics, politicians, and civil society groups. The digital revolution has significantly affected the constitutional ecosystem. Private multinational companies producing, managing, and selling digital products and services have emerged as new dominant actors alongside nation states. Using digital technology, both states and big technology companies restrict individual

* Assistant Professor of Law, Technology and Innovation, and Programme Director of the European Master in Law, Data and AI (EMILDAI), School of Law and Government, Dublin City University, Ireland. Email: edoardo.celeste@dcu.ie.
fundamental rights in multifarious ways. A series of intrinsically “constitutional” questions emerge on how to generalise and re-specify the key principles of contemporary constitutionalism in the novel context of the digital society. Scholars have advocated the need for a digital constitution, including the idea of drafting an “internet bill of rights,” a written document on the model of the ancient declarations of rights of the eighteenth century. Following this appeal, many civil society groups have crafted their own decalogue of digital rights, giving rise to a movement advocating a new form of “digital constitutionalism.”

This paper investigates whether these internet bills of rights add value in the constitutional ecosystem. In particular, this paper reconstructs the genealogy of the idea of an internet bill of rights among the scholarly models advocated for a digital constitution (I), and illustrates the extent to which the declarations of rights emerged in the last few years can play a compensatory and stimulatory role in the current process of constitutionalisation of the digital society (II). This paper will posit that these declarations will highlight the transformative potential of societal constitutionalism. By adopting the typical language of constitutions, they seek to be part of the current conversation on how to translate the core values of contemporary constitutionalism in a digital society. However, in contrast to other constitutional instruments, internet bills of rights are not legally binding sources nor the output of institutionalised processes of deliberation. Thus, they appear as a more ductile instrument whereby their promoters are free to experiment with new legal solutions in a gradual way, through a multitude of initiatives, and in a more democratic manner. This includes actors beyond the worlds of politics and business.

II. FROM THE DECLARATION OF INDEPENDENCE OF CYBERSPACE TO THE DIGITAL CONSTITUTION

Early internet literati considered cyberspace as a recently discovered, uninhabited island and a place where a new beginning was possible. In 1994, Esther Dyson, George Gilder, George Keyworth, and Alvin Toffler published an article solemnly titled “Cyberspace and the

American Dream: A Magna Carta for the Knowledge Age."³ Cyberspace was celebrated as the “land of knowledge” whose exploration would represent “civilization’s truest, highest calling.”⁴ The border between the physical and virtual worlds was considered as the “electronic frontier” delimiting a new prosperous Far West.⁵ In 1996, John Perry Barlow wrote the famous “Declaration of the Independence of Cyberspace.”⁶ This text reiterated the idea of cyberspace as the “home of mind,” a “global social space” where anyone “may express his or her beliefs, no matter how singular, without fear of being coerced into silence or conformity.”⁷ As its title clearly evokes, the aim of the declaration was to assert the independence of cyberspace from the physical jurisdiction of nation-states, defined as “weary giants of flesh and steel,” on the basis of its “natural” autarky.⁸

This worldview, subsequently dubbed “cyberlibertarianism,” posited that cyberspace was a “speech-dominated” world where no physical harm was conceivable.⁹ This assumption would have justified the argument of internet exceptionalism, (i.e., that nation-states do not have the right to apply their rules to cyberspace; conversely, the digital world could find its very own system of governance).¹⁰

A second group of scholars, at their turn dubbed “cyberpaternalists,” strongly rejected this vision of cyberspace as an intrinsically distinct world of pure speech, capable of self-regulation.¹¹ They argued that purely cyberlibertarian and cyberanarchist views were utopian: these ideals would be destined to surrender to the supremacy of the feral law of the internet environment. In cyberspace, the “invisible hand” would not be guided by democratic and egalitarian archetypes but by the

---


4. Id.

5. Id.


7. Id.

8. Id.


11. See MURRAY, supra note 2; see, e.g., Jack L. Goldsmith, *Against Cyberanarchy*, 65 UNIV. CHI. L. REV. 1199 (1998); see, e.g., LAWRENCE LESSIG, *CODE: AND OTHER LAWS OF CYBERSPACE, VERSION 2.0* (2006); see, e.g., Boyle, supra note 9; see, e.g., Joel R Reidenberg, *Yahoo and Democracy on the Internet*, 42 JURIMETRICS 251 (2002).
dominant pair of government and commerce.\textsuperscript{12} The internet would degenerate from the ideal “home of mind” into a space of commercial profit and government surveillance.\textsuperscript{13} For this reason, cyberspace cannot be considered as an autonomous space, fully distinct from the physical world, and should be subject to regulation.\textsuperscript{14} As stated by the German Federal Constitutional Court, “even the Internet cannot create a legal vacuum.”\textsuperscript{15}

Unfortunately, over the past few years, a series of events has inexorably infringed two central tenets of the cyberlibertarian vision. It suffices to recall the mass surveillance programmes unveiled by Edward Snowden in 2013 and the recent Cambridge Analytica scandal involving Facebook, to understand that cyberspace is not the utopian, self-regulating realm of the mind advocated by the cyberlibertarians.\textsuperscript{16} Moreover, today it is apparent that cyberspace is not a remote island: it is a distinct, autonomous, and watertight environment separable from the physical reality theorised by those early internet thinkers. Recent technological developments have contributed to bind physical and virtual existence together in a way that, today, for many, these two are complementary and inseparable. The use of digital technology has become an integral part of our daily life, even more after the recent COVID-19 pandemic.\textsuperscript{17} Reading news, communicating, searching for a job, and professing one’s own political or religious faith are all examples of activities that many individuals habitually perform online. Today, digital technology does not merely represent one of the many instruments available to exercise a broad range of fundamental rights rotating around the exchange of information, but has even become one of the most important catalysts of these rights.\textsuperscript{18}

The internet environment is not a \textit{terra nullius}, and does not lie anymore at the peripheries of our lives. It is a context where nation-states will exercise their sovereignty, albeit in a complex way, due to the

\begin{itemize}
\item \textsuperscript{12}\textsuperscript{12} LESSIG, supra note 11 at 4.
\item \textsuperscript{13} See LESSIG, supra note 11; see also Boyle, supra note 9.
\item \textsuperscript{14} See MURRAY, supra note 2.
\item \textsuperscript{15} Vorratsdatenspeicherung [Data Retention], 260 (Ger.), https://germanlawarchive.uscomp.org/?p=1200.
\item \textsuperscript{17} See TERO Karppi, Disconnect (2018); Gilles Dowek, Vivre Aimer, Voter En Ligne Et Autres Chroniques Numériques (2017); see also Sherry Turkle, \textit{Life on the Screen} (1995); see also Annette N. Markham, \textit{Life Online: Researching Real Experiences in Virtual Space} (1998).
\item \textsuperscript{18} See Packingham v. N.C., 137 S. Ct. 1730, 1734 (2017).
\end{itemize}
emergence of other dominant transnational private actors. It is an ecosystem that plays a central role in determining the extent to which we can exercise a broad array of our fundamental freedoms. On the one hand, the virtual world has increased the standard to which we are accustomed to exercise our rights. Digital technology offers unprecedented opportunities that are not equated by other existing media. On the other hand, contrary to what the cyberlibertarians argued, it is not true that “bytes can never hurt me.” Cyberspace can indeed be an instrument to perpetrate crimes with direct repercussions in the physical world.

The internet environment has not only acquired legal relevance, but also an apparent constitutional relevance. Protecting fundamental rights and regulating the exercise of power in the digital environment have become major challenges of contemporary constitutionalism. Everyone should have the right to exercise their rights according to the standard that new technology allows and be protected against potential fundamental rights infringements. This need is informing a multilevel process of constitutionalisation involving multiple normative counteractions, ranging from the adoption of more traditional constitutional instruments to more innovative tools emerging even beyond the state dimension. This paper will focus on one peculiar solution: the appeal to adopt a digital constitution. An ambitious demand, as we will explore in the next few sections, which was originally advocated at the academic level following different models.

A. Lessig: The Constitution in the Internet

In his seminal book Code, Lawrence Lessig highlights the role of the internet architecture as a powerful and effective regulatory
mechanism. Technology companies would have the ability to directly shape the code of software determining how users can interact in the digital world. Joel Reidenberg had already analysed this form of private rulemaking through technical solutions under the name of lex informatica. Lessig uses the slogan “code is law.” The legal norm somehow becomes invisible, embedded in the structure of the internet. Internet users are no longer offered the possibility to infringe the law because respecting it is the only action still available to them.

The code, as a form of physical constraint, nullifies the space between law and law enforcement. Code is law in action and, as highlighted by Lessig, could also be a constitution in action. The internet architecture enshrines constitutional values, in principle those determined by technology companies. Code would represent the internet’s material constitution, its factual constitutional order.

However, Lessig rejects the idea of code’s determinism. The internet architecture should not remain that which is shaped by technology companies, completely subject to their economic interests. Code—he argues—would be malleable. He warns against the “fallacy of ‘is-ism.’” One should not think that how the internet is now corresponds to how it should be. In contrast to the views of cyberlibertarianism, Lessig mistrusts the capability of cyberspace to self-regulate. In order to protect fundamental rights in the digital environment, he proposes to correct code by incorporating the values of the constitution, in his case, the US Constitution. Rather than praising the existing constitution of the internet, Lessig advocates a constitution in the internet.

B. Teubner: Civil Constitutions

Lessig wisely recognised that instilling the constitution in the code is not a straightforward task. It implies a complex work of “translating” constitutional values, a problematic process of selection, interpretation, and transformation of principles entrusted both to the legislator and to the judiciary. The issue of translation of traditional constitutional

24. See LESSIG, supra note 11.
26. See LESSIG, supra note 11.
27. See id. at 4.
28. Id. at 25.
29. Id. at 31
30. Id. at 4.
32. LESSIG, supra note 11, at 157, 315, and 325.
values in the context of the digital society is also pivotal in the works of Gunther Teubner. The German scholar generally questions the effectiveness of national constitutional law vis-à-vis three major trends of the twenty-first century: digitisation, privatization, and globalisation. The challenge of constitutional theory would be to “generalise its nation-state tradition in contemporary terms and re-specify it.” This operation, according to Teubner, would necessarily imply to disconnect constitutional law from the state dimension and eventually theorise a constitutionalism without the state.

Teubner, therefore, moves a step forward in comparison with Lessig. He shares the view that the lex electronica, what in the previous section was called the internet’s material constitution, is “corrupted” by the economic interests of technology companies. The lex electronica is for Teubner the heir of the lex mercatoria in the context of the cyberspace. He argues that “it is here that the constitutional question of the Internet arises.” The set of constitutional rules self-produced by dominant societal actors, such as technology companies, would suffer from a structural bias. They would only be of a “constitutive,” and not “limitative,” nature. This structural corruption would in turn trigger a reaction from what Teubner calls the “spontaneous sectors” of society, such as governmental agencies, civil society groups, trade unions, consumer protection organisations, and the like. The interaction between dominant players and this category of actors would engender a process of constitutionalisation and, ultimately, the emergence of new constitutional norms.

Given the transnational nature of the digital society, for Teubner, the digital constitution cannot derive from classical political processes of constitutionalisation. The alternative he proposes is a “societal constitutionalism,” in which subsectors of society would autonomously produce constitutional norms that would subsequently be absorbed by the legal system in a process of mutual influence. Therefore, in

34. Id. at 4.
35. Id. at 21.
36. Id. at 20.
38. See TEUBNER, supra note 37, at 94.
39. See TEUBNER, supra note 37.
contrast to Lessig, Teubner does not advocate a direct translation of constitutional principles drawn from national constitution into the internet code, but rather theorises a complementary constitutional process: the emergence of a multiplicity of “civil constitutions.”

Even if one could argue that for Teubner the ultimate mission of a digital constitution remains the same—the protection of fundamental rights in the digital society—its intrinsic characteristics would be very different from Lessig’s archetype of constitution in the internet. Teubner’s digital constitution emerges beyond the state. It is potentially plural, resulting from a multiplicity of societal interactions. It is civil, in the sense that it is issued by the civil society, and it is not the result of institutionalised political processes. It is crowd-sourced, since its elaboration involves a variety of different actors. And, lastly, it is characterised by an unconventional nature, being something between a “legal text” and a “de facto structure of social systems.”40 All peculiarities that we will interestingly find again in the phenomenon of the emergence of internet bills of rights, which will be examined in the second part of this paper.

C. Pernice: The Constitution of the Internet Society

Pernice shares one of the fundamental premises of Teubner’s thought: the insufficiency of state-centred constitutionalism to address the challenges of contemporary society. A series of global issues, such as international terrorism, climate change, global financial regulation, internet governance, and world peace-keeping, cannot be efficiently addressed at a national or regional level.41 Policies concerning these problems, when developed at local levels, unavoidably produce external effects, structurally lack democratic legitimacy, and are capable of undermining people’s self-determination.42 The power architecture no longer exclusively pivots on the states, but is more similar to a “constellation” of dominant actors, including private entities acting across a multiplicity of national territories.43 In order to preserve our fundamental constitutional values in this mutated societal scenario, a

40. TEUBNER, supra note 37, at 19.
42. See id.
global model of constitutional governance is needed. Pernice analyses existing paradigms of regulation beyond the state, such as the EU, the UN, the WTO, the so-called Rio Process on sustainable development, and internet governance, concluding that none of them, taken singularly, provide a universal remedy for all the challenges of contemporary society. For Pernice, the solution consists in a form of multilevel global constitutionalism, a complex architecture where all these partial legal systems would complement each other in order to shape a “global constitutional framework.” In such a setting, state constitutionalism is not absent. In contrast to Teubner’s vision, it crucially lies at the basis of this fragmented and pluralistic construction, and, to this end, it is necessarily rethought. Pernice interprets the constitution in a “postnational” sense, no longer uniquely anchored to the state, but opened up to the participation of individuals, recognising the important co-regulatory role of private actors.

Pernice’s recent works highlight the twofold role that the internet plays in this scenario. The digital revolution is seen both as one of the transformative elements of contemporary society and as a unique opportunity to achieve a global constitutional framework. The internet is not only a source of threat to our fundamental rights, but it is also pivotal in allowing individuals to share information and participate in democratic life and in enhancing transparency within the polity. For Pernice, the current impact of the digital revolution is so significant that he denotes contemporary society as the “Internet society.” Digital technologies create a “global community” of individuals and shape a “global public sphere” of political interaction. In this way, the internet becomes the driver of the new global constitutional framework, what Pernice also calls the constitution of the Internet society.

However, in this context, the internet is not only the catalyst to a form of multilevel global constitutionalism, but, as said before, it is also the object of a global model of governance. Pernice, therefore, adopts the expression “constitution of the Internet” to specifically denote the basic legal order of internet governance. Such a constitution would emerge

44. Pernice, supra note 41.
45. See id.
47. PERNICE, supra note 43.
48. Id.
49. Pernice, supra note 41, at 6.
50. Pernice, supra note 41.
51. Id.
through an autonomous process of self-constitutionalisation carried out by a plurality of actors beyond the state.\textsuperscript{52} The emerging constitution of the internet would represent one of several partial constitutions composing the multilevel mosaic of the global constitutional framework. In contrast to the broader notion of the “constitution for the Internet society,” its mission would therefore be to preserve the open, free, and interoperable character of the internet, its variegated institutional framework, and the multi-stakeholder character of the processes of deliberation characterising its governance.

According to Pernice, the constitution of the internet will not be realistically enshrined in an international treaty.\textsuperscript{53} As with Lessig and Teubner, it could embody a basic legal order without being expressed in a written text.\textsuperscript{54} Echoing Teubner, Pernice argues that the constitution of the internet could remain fragmentary, thus very dissimilar from the traditional idea of a constitution.\textsuperscript{55} Nevertheless, these characteristics do not diminish the role of the internet constitution in a broader context. In contrast to Teubner, Pernice offers a wider vision. The constitution of internet governance can provide useful examples to understand how to craft constitutional norms for the global society.\textsuperscript{56} He considers the constitution of the internet as one of the many “interwoven partial constitutions” whose elements, be they principles or institutional mechanisms, can be successfully harnessed to build the global constitutional framework of the internet society.\textsuperscript{57}

\textbf{D. Rodotà: A Charter of Rights for the Internet}

While Teubner theorises the progressive emergence of a digital constitution from the societal substratum through a complex process of interactions occurring outside traditional political processes, Stefano Rodotà directly advocates a cooperation of all the relevant stakeholders to draft a digital constitution.\textsuperscript{58}

In 2005, the Italian legal scholar first wrote about the need to draft

\textsuperscript{52} Id. at 22.
\textsuperscript{54} See id. at 23.
\textsuperscript{55} Id.
\textsuperscript{56} See PERNICE, supra note 43.
\textsuperscript{57} See PERNICE, supra note 43, at 26.
\textsuperscript{58} See Stefano Rodotà, \textit{Una costituzione per Internet?} [A Constitution for the Internet], POLITICA DEL DIRITTO 337 (2010).
a “Charter of Rights” to harness the development of the internet.59 Both states and powerful corporations would threaten fundamental rights online. The internet would become an instrument to control, monetise, and discriminate individuals to reduce their specificity and autonomy. As the Parisians (re)discovered the importance of Mona Lisa when the famous Da Vinci painting was stolen from the Louvre museum in 1911, so too would these series of new threats allow us to perceive the lack of guarantees for the digital society.60 Therefore, at the moment, what would be most needed for Rodotà is a series of guiding principles. In his words:

Time is come to affirm the principles of the new global people: freedom of access, freedom of use, right to knowledge, right to privacy, protection of the commons. Time is come to recognise these principles in a new Charter of Rights.61

Rodotà does not believe in a spontaneous natural reaction of the internet. Invoking the original libertarian nature of cyberspace to reject the possibility to craft a charter of rights for the internet would represent a “position of rearguard”—writes Rodotà in another article.62 It would be short-sighted not to take into account the multifarious situations in which, today, our fundamental rights are restricted in the online sphere by both nation-states and private corporations, as in the early cyberlibertarian charters. For the Italian scholar, elaborating a charter of rights for the internet does not mean caging the freedom of the digital environment; conversely, charters and declarations have traditionally been the basis for an expansion of rights.63 The formal recognition of rights in a written document would never be the end of the story—the (written) rights are instruments to fight for the (real) rights.64

60. See STEFANO RODOTÀ, IL DIRITTO DI AVERE DIRITTI 16 (2012); (drawing from Dominique Rousseas’s example in, La démocratie ou le vol de La Joconde, in NOUVELLES QUESTIONS SUR LA DÉMOCRATIE 145 (Bertrand Mathieu et al. eds., 2010)).
61. Rodotà, supra note 58 (specifically, the author’s translation).
63. Id.
64. RODOTÀ, supra note 60, at 32.
Rodotà therefore abandons the innovative schemes proposed by Lessig, Teubner, and Pernice to stress the crucial role that a classical constitutional instrument, such as a charter of rights, could play in the digital society. For Rodotà, the idea of a digital constitution reacquires its traditional meaning of a written document—it neither refers to the code nor to a constitution as is intended in the common law systems—as normative “architecture” resulting from the combination of various legal sources.  

Declarations of rights are texts that have marked epochal revolutions in the history of human rights. Rodotà rediscover their power. He decides to exploit the evocative value that they have progressively acquired over the past few centuries to convey the new message of digital rights. In an article written in English and published in 2008, Rodotà employs for the first time the term “Internet Bill of Rights,” and explains: “The choice of the old formula of the Bill of Rights has a symbolic force, it underlines that the aim is not to restrict freedom on the web but, on the contrary, to maintain the conditions for letting it continue to prosper. To do so, ‘constitutional’ guarantees are required.”

The identification of a suitable instrument to promote fundamental rights in the digital environment is central for Rodotà. Like Lessig, Teubner, and Pernice, Rodotà understands that the current major challenges of the online world directly affect the constitutional dimension of the protection of individual rights. Therefore, Rodotà’s solution seems to be the most conventional: he proposes to use a classical constitutional instrument to solve a constitutional problem. Nevertheless, the idea of a charter of rights for the internet, even though it appears at first sight in line with the constitutional tradition, presents some original aspects.

Rodotà clarifies that the use of the descriptors “charter” or “bill of rights” does not imply that the digital constitution should be the output of elaborate processes similar to those that have brought about the adoption of those documents in the past. For the Italian scholar, the charter of rights for the internet should not be the result of top-down processes analogous to those that led to the so-called “constitutions octroyées.” Nor to the charters graciously granted by enlightened

---


67. Cf. Gaetano Azzariti, Internet e costituzione 367 (2011) (the author argues that using the descriptor “internet bill of rights” is legitimate in so far as it evokes the symbolic values of this concept, but cannot lead to think that such a document has the same qualities of traditional constitutions).
princes in the nineteenth century, or to the constitutions deliberated by constituent assemblies in the twentieth century. Sharing Pernice’s scepticism, Rodotà contends that the digital constitution cannot derive from traditional processes of international lawmaking, which would require the participation of all the relevant stakeholders: therefore not only states, but also corporations, and—especially—citizens. Only in this way could the digital constitution reflect the participatory and open nature of the internet, “the largest public space that the humankind has ever known.”

Consequently, for Rodotà, the charter of rights for the internet should be necessarily global. In his book “Il diritto di avere diritti” (The right to have rights) he wrote:

Landless rights wander in the global world in search of a constitutionalism that, being it global as well, could offer them a safe mooring. Today, rights are orphans of a land where they had their roots: national sovereignty once offered them a solid protection. Today, rights are dissolved in a world without borders, dominated by unrestrained power.

Rodotà reports the legend of a man who owned a mill in Sanssouci, very close to the royal residence of Frederick the Great. The German king, disturbed by the proximity of the building, threatened to seize the miller’s property using his royal prerogatives. However, the humble man politely reminded the king of the existence of a high court in Berlin, which had the power to assess the justness of the royal order. Rodotà draws a comparison between the current global context in which rights and individuals have lost their reference points, and wonders “who is the king and where are the judges today?”

For Rodotà, we now live in a global society where we do not know what our guiding principles are and, even less, where we can obtain judicial redress. He argues that a global charter of rights for the internet presupposes a rethinking of the relationship between public

68. Rodotà, supra note 59.
69. Id.; see also Rodotà, supra note 58 (Rodotà in particular praises the model of the “dynamic coalitions” of the Internet Governance Forum); Cf. PERNICE, supra note 43.
70. Rodotà, supra note 58 (use author’s translation).
71. RODOTÀ, supra note 60, at 3 (use author’s translation).
73. RODOTÀ, supra note 60, at 3 (use author’s translation).
and private. He favours an overtaking of strategies of pure “vertical domestication,” where supra-national norms are in turn incorporated and implemented at lower levels. He advocates a “horizontal,” expanding construction of norms, stressing in particular the future role that judges will have to interpret the general principles of the charter.74

E. Following Rodotà’s Model

Rodotà’s idea of a charter of rights for the internet attracted a significant following among the scholarship. The model of an internet bill of rights was considered as a suitable solution to limit the powers of predominant players in the digital environment, and to protect the fundamental rights of individuals, in particular by updating or re-contextualising existing principles.75 A few scholars further stressed that this instrument should also create new institutions.76

Giovanna De Minico considered the question of which “legislative body” or “constituent power” should draft the internet bill of rights.77 After having rejected the idea of a charter drafted by one single state or by a coalition of states, she concluded that the best solution would be to create a “public supranational authoritative body.” This entity, according to De Minico, should legislate in a way that resembles the American “notice and comment” procedure (i.e., by preliminary gathering the opinions of all the stakeholders involved before making its final decision).

Kinfe Micheal Yilma made a case for a declaration of internet rights adopted by the UN General Assembly.78 The structure of the existing UN organs would suitably fit the process of elaboration of a similar document. In particular, he praised the potential input the General Assembly could receive from existing crowd-sourced mechanisms.

74. Rodotà, supra note 58.
76. See Casacuberta, Senges, and Duart, supra note 75; Musiani, supra note 75; De Minico, supra note 75.
77. De Minico, supra note 75.
78. Yilma, supra note 75.
headed by the Office of the High Commissioner for Human Rights in the phase of drafting. Furthermore, a similar declaration would have a strong normative impact, while at the same time avoiding hard legislation.

However, in the scholarship there is no consensus on the role of states in drafting the charter of rights for the internet. De Minico and Yilma, while recognising the global dimension of the digital environment, entrusted a central role to states and intergovernmental organisations. Conversely, Andreas Fischer-Lescano, echoing Teubner, stated that the peculiarities of the digital environment impose a constitutional solution that necessarily overtakes the traditional nation-state dimension.79

Lastly, it is interesting to report Gaetano Azzariti’s development of the idea of a charter of rights for the internet.80 Azzariti recognised that an effective response to the issue of the protection of fundamental rights on the internet should have a global nature. Nevertheless, he argued that a digital constitution should not necessarily be conceived as a single, “cosmopolitan” document addressing all of humankind. This would amount to a deplorable form of “constitutional colonialism.”81 For Azzariti, the route to follow is longer and more complex. He supports the slow process of elaboration of a multiplicity of charters and declarations by a variety of actors, including the civil society. This multitude of charters would persuade, influence, and create a dialogue. Only in this way, the definition of the fundamental rights for the digital society could be achieved in a democratic and participatory way. Though criticising Teubner, Azzariti eventually theorises a pluralisation of the idea and the process of elaborating an internet constitution, somehow echoing the archetype of civil constitutions proposed by the German scholar and merging it with Rodotà’s model of digital constitution.

III. THE ROLE OF INTERNET BILLS OF RIGHTS

A. Digital Constitutionalism

Azzariti’s idea of a multiplicity of charters of rights for the internet emerging simultaneously is not a purely academic proposal, or a utopian theoretical abstraction.82 When suggesting plural processes of

---

79. Fischer-Lescano, supra note 75; cf. LESSIG, supra note 11 (who opts for a residual role of states, which would be called to intervene only when the risk of adverse public consequences emerges); see also Teubner, supra note 33.
80. Azzariti, supra note 67.
81. See id.
82. See id.
elaborating fundamental principles for the digital society, he probably drew his inspiration from reality. The idea of a written charter of rights for the internet, indeed, did not remain a theoretical prototype, subject of discussion for academics and internet literati. It was translated into practice, and interestingly, we can identify only one or few single texts that attempted to translate this idea of an internet constitution. We observe a significant phenomenon, the emergence of numerous charters of rights for the internet, very similarly to what Azzariti preconized: what we could define as a *movement* of “digital constitutionalism” characterizing an intense constitutional *moment.*

Over the past few years, many civil society groups have published their proposal for a charter of digital rights. To mention some recent examples, in 2014, the African Declaration Group issued the “African Declaration on Internet Rights and Freedoms” and the Just Net Coalition promoted the “Delhi Declaration for a Just and Equitable Internet.” In 2015, Rodotà himself chaired a commission composed of politicians, academics, and representatives of private companies and NGOs, and issued a “Declaration of Internet Rights,” subsequently endorsed by the Italian Chamber of Deputies. In 2016, a German working group supported by the Zeit-Stiftung proposed a “Charter of Digital Fundamental Rights of the European Union.”

A significant number of individuals, including academics, politicians, and technology experts, also proposed their own draft of an

83. See CELESTE, supra note 1; Edoardo Celeste, *Digital constitutionalism: a new systematic theorisation*, 33 INTERNATIONAL REV. L. COMPUTERS & TECH. 76 (2019); Celeste, supra note 19; Dennis Redeker, Lex Gill & Urs Gasser, *Towards digital constitutionalism? Mapping attempts to craft an Internet Bill of Rights*, 80 INT’L COMMC’N GAZETTE 302 (2018) (in my previous works, I defined “digital constitutionalism” as an ideology informing an ongoing process of constitutionalisation that is seeking to adapt the core values of contemporary constitutionalism to address the challenges of the digital age. Such a process is of such relevance and scale that I spoke of a “constitutional moment.” At the same time, scholarship from cognate disciplines, such as political science and communication studies, have highlighted the impact of the initiatives leading to adoption of charter of rights for the Internet as a group: hence I propose here to speak also of a movement of digital constitutionalism).


internet bill of rights. Robert Gelman published his “Declaration of Human Rights in Cyberspace” back in 1997.88 In 2010, Andrew Murray advanced “A Bill of Rights for the Internet.”89 In 2015, Mike Godwin wrote “The Great Charter for Cambodian Internet Freedom.”90 Interestingly, there have been discussions at the national level too on the need to incorporate a charter of internet rights into domestic legislation.91 In 2014, the Brazilian parliament was the first to approve a statute establishing “principles, guarantees, rights and obligations for the use of Internet.”92 In 2019, the Nigerian national assembly voted in favour of a “Digital Rights and Freedom Bill,” which was, however, not signed into law by the president.93 Even international organisations have issued a series of declarations proclaiming rights and principles for the digital society. The EU Commission has recently announced its intention to promote the adoption of a solemn declaration on digital rights and principles.94 One can also mention the Recommendation of the Committee of Ministers of the Council of Europe on Internet freedom of 2016,95 and the Code of Ethics for the Information Society,

91. GREEN PARTY (NEW ZEALAND), Internet Rights and Freedoms Bill (2014); LIBERAL DEMOCRATS, Creating a 'Digital Bill of Rights': Why do we need it, and what should we include?, (2015), https://d3n8a8pro7vhmx.cloudfront.net/libdems/pages/8730/attachments/original/1428513286/Digital_Bill_of_Rights_Consultation_Paper_FINAL.pdf?1428513286.
adopted by the UNESCO in 2011.\textsuperscript{96}

As one can notice from the examples mentioned above, these documents have different names. Mass media generally refer to these initiatives with a variety of denominations, such as “Internet bill of rights,” “online Magna Carta,” or “Internet constitution.”\textsuperscript{97} This article reflects the use established by some scholars to adopt the descriptor the “internet bills of rights.”

B. Generalisation and Respecification

The emergence of a multiplicity of internet bills of rights does not represent an isolated phenomenon in the contemporary constitutional landscape. We are currently witnessing a new constitutional moment. The digital revolution generates a series of changes at a societal level, which ultimately ferment under a vault of constitutional norms shaped for an analogue society. The constitutional ecosystem is gradually reacting to face the challenges created by the advent of digital technology. The core values of contemporary constitutionalism are progressively translated in the context of the digital society. However, the constitutional discourse is no longer uniform. It resembles a normative conglomerate. It is fragmented, and acts on multiple, intersecting levels. It involves public and private, dominant and weak actors. Constitutional norms flow beyond the state dimension to reflect societal dynamics occurring in a transnational and virtual space.

This paper argues that internet bills of rights are an integral part of the constitutional response to the challenges of digital technology. The emergence of these texts is a component of the process of constitutionalisation that, at multiple levels, is adapting the main principles and values of contemporary constitutionalism to the needs of the digital society. However, their role in such a dynamic is generally


neglected, and the reason for this disregard can be easily explained by looking at the legal value of these texts. Despite the constitutional “tone” that they formally adopt, not only do they not have any binding legal value from a formal perspective but they can hardly be regarded as presently exercising any legal influence from a substantial point of view.98

Most of the scholarship exclusively saw the value of internet bills of rights in their future legal transformation, codification, or incorporation into a (more) binding legal source. Some authors, for example, emphasise the role that international institutions, such as the United Nations, could play in this regard by formally including an internet bill of rights within their body of rules.99 Interestingly, Luca Belli, commenting on the Italian Declaration of Internet Rights, proposed to integrate it within the Italian Constitution as a separate instrument, taking as a prototype the French “Charte de l’environnement,” a document enshrining a series of rights and principles related to the protection of the environment that was formally incorporated into French constitutional law in 2005.100 Only a few scholars highlighted their utility as an intermediary step to enhance consensus and promote discussion in view of the adoption of an international normative framework on the topic or, more generally, as an instrument with a “cultural and political” value, yet always underlining their lack of binding legal force.101


99. See Luca Belli, Dichiarazione dei diritti in Internet. Cuius Regio eius Religio?, in Verso Un Internet Bill of Rights 37 (2015); see also Monica Alessia Senor, La forma della Dichiarazione dei diritti in Internet, in Verso Un Intenet Bill of Rights 69 (2015); Yilma, supra note 75; cf. De Minico, supra note 75 (provides a theoretical perspective, suggests the idea of a supranational legislator).


101. See Nicolò Zingales, Mettiamo la Dichiarazione dei Diritti in Internet in prospettiva, in Verso Un Internet Bill of Rights 73–79 (2015) (comparing the Italian Declaration of Internet Rights with atypical sources of international law with exhortatory value, such as the reports of the UN Special Rapporteurs); see Marco Bassini & Oreste Pollicino, Carta dei diritti in Internet verso una missione culturale. Né costituzione né legge., in Verso Un Internet Bill of Rights 129–134 (2015) (stressing the “cultural and political value” of the Italian Declaration of Internet Rights in “guiding” the future choices of the Italian legislator); see also Claudia Padovani & Mauro Santaniello, Digital constitutionalism: Fundamental rights and power limitation in the Internet eco-system, 80 Int’l Commc’n Gazette 295–301 (2018) (highlighting the political mission of Internet bills of rights).
It is true, realistically speaking, internet bills of rights might hardly be directly elevated in the future to the rank of soft or hard law by the judiciary or the legislature but could certainly be used as a source of inspiration for judicial decisions and legislative initiatives. Internet bills of rights do not generally have a legally binding value at present but possess a legal status as they play a constitutional role. They contribute to build and shape the narrative that forms the societal background where legal actors draw inspiration to interpret the analogue nomos that is currently struggling to address the challenges of the digital revolution and help make it evolve.\textsuperscript{102}

It is precisely in their present, non-legally binding value—this paper argues—that the “force” of internet bills of rights lies.\textsuperscript{103} It is their flexibility and ductility that enhance their capacity to experiment and represent the avant-garde of the constitutional discourse that is translating our core fundamental rights and principles in the new context of the digital society. And this translation, building on Teubner’s theory of societal constitutionalism, consists in a twofold process of “generalisation” of the principles of contemporary constitutionalism followed by their “re-specification” to specifically address the challenges of the digital society.\textsuperscript{104}

Internet bills of rights can, therefore, be seen as proto-constitutional discourses because, by linking society and law, they translate societal needs in the language of constitutional norms and transmit such impulses to the other components of the constitutional dimension.\textsuperscript{105} They are proto-constitutional because they advance the elaboration of constitutional principles for the digital society, and, as a part of a porous constitutional conglomerate of norms, they allow new formulations of values and ideals to flow between constitutional layers, as between communicating vessels.

In this way, in the following sections, the same characteristics of internet bills of rights that have often been criticised, such as the fact that they lack binding legal value because they are issued from civil society, are non-institutionalised processes, adopt a constitutional tone without having any formal constitutional value within national legal systems, and their significant number, will be read as a source of force


\textsuperscript{104.} See TEUBNER, supra note 37; Celeste, supra note 19.

\textsuperscript{105.} Cf. TEUBNER, supra note 37, at 94 (describing the process of “self-constitutionalisation of corporations” as a “translation process” of supra-national soft law into the normative dynamics of corporations).
INTERNET BILLS OF RIGHTS

of these declarations.106

1. Experimentalism

From a cursory look at the examples of internet bills of rights listed above, it is easily detectable that most of them are not issued from institutional processes, but are rather the output of civil society organisations or even single individuals. Even when they are adopted in the form of a recommendation or declaration by international institutions, such as the OECD or the Council of Europe, these documents do not have any binding legal value. At first sight, internet bills of rights appear as a weak instrument within the mosaic of norms that address constitutional issues in the digital society. There is no doubt that, if contrasted to the internal rules of private actors or to state law, for instance, these documents cannot bear the comparison in terms of enforceability.

However, this view takes the legal force as a unique parameter to assess the role of these documents within the constitutional ecosystem. In this way, one would neglect the very reason as to why internet bills of rights are emerging. We have seen that these documents are not the only ones seeking to translate our fundamental values and principles to address the challenges of digital technology. Internet bills of rights are part of a conglomerate of norms emerging at multiple levels. Therefore, it is not unreasonable to posit that their peculiarity—their lack of legal force—is also one of the main factors at the basis of their emergence.

In contrast to the other normative instruments which are currently conveying constitutional counteractions against the challenges of digital technology, internet bills of rights, being non-legally binding, appear to be particularly ductile, plastic, and malleable. These texts are unaffected by the formalisms and constraints that would ensue from an institutional adoption. Internet bills of rights, as other normative instruments are currently doing, aim to translate our constitutional values and principles in the context of the digital society, but, in contrast to them, they are freer to experiment and innovate.107

For example, these texts enshrine innovative rights and principles

106. See Verso un Internet Bill of Rights, supra note 96 (providing a critical account on the Internet Bill of Rights); see also Yilma, supra note 75.

on which there is no consensus and that have not been incorporated into other constitutional instruments yet. In this way, internet bills of rights are advancing the discourse of constitutionalism in the digital age. Like the declarations of rights of the eighteenth century, they are proclaiming new principles, a necessary step to nourish the conversation on their content and shape, which, in the longer term, will lead to their factual enforcement. Indeed, as Rodotà wrote, even the mere action of declaring rights is a first move towards their substantial recognition.108

Beyond enhancing their experimentalist character, the lack of binding legal force of internet bills of rights also allows a higher level of plasticity in their processes of elaboration and deliberation. Internet bills of rights are the output of innovative procedures involving a plurality of actors. Not only traditional institutional players, but also representatives from the technology industry, NGOs, civil society groups, and single individuals. Advancing the discourse of constitutionalism in the digital age through non-legally binding instruments allows for the adoption of more inclusive processes of deliberation, which may not be offered through traditional political channels. The procedures which lead to internet bills of rights are tailored to reflect a global, transnational society in which, as seen in the first part of this work, private corporations managing the access to digital technology have emerged as new dominant actors beside states.

Internet bills of rights offer a way to convey the instances of what Teubner calls the “spontaneous” sphere of the society, such as NGOs and advocacy groups, and, as we will see in the next section, to put them in dialogue with the powerful actors of politics and the economy.109 The elaboration of these documents recreates a political dimension, “le politique,” in a global context, where the traditional political world, “la politique,” anchored to the state dimension, could no longer be of help.110 Internet bills of rights bring the conversation on constitutionalism in the digital age closer to the civil society. Bespoke and innovative processes of deliberation enhance the proximity of this discussion to the people and ensure greater inclusiveness. In this way, these documents

---

108. Rodotà, supra note 60, at 32, 75.
109. See Teubner, supra note 37, at 91; see also Christoph B. Graber, Bottom-up Constitutionalism: The Case of Net Neutrality, 7 TRANSNT'L L. THEORY 524, 546 (2016).
110. See generally Teubner, supra note 37, at 114 (providing an analysis of the dichotomy “la politique versus le politique” in the context of societal constitutionalism); see also, Oliver Marchart, Post-Foundational Political Thought: Political Difference in Nancy, Lefort, Badiou and Laclau (2007); cf. Niklas Luhmann, Law as a Social System 487 (2004), (author provides an argument that “the structural coupling of the political system and the legal system through constitutions does not have an equivalent at the level of global society”); see also Graber, supra note 109, at 529 (providing a commentary on Luhmann’s position).
emerge as a counteraction to the progressive phenomenon of privatisation of the law in the digital society. Internet bills of rights ultimately aim to foster democracy in a context dominated by the *lex digitalis* of powerful multinational corporations and by a declining, distant politics.

2. Communicability

Internet bills of rights explicitly evoke the constitutional dimension by adopting a constitutional tone. The denominations of these texts refer to traditional constitutional instruments: the charters and declarations of rights. They reproduce the structure of constitutional texts; they employ their typical jargon and rhetoric. Yet, as we have seen, internet bills of rights do not exercise any legal force. They do not have the typical superior value of constitutional texts within the hierarchy of legal sources.

However, once again, in line with what has been argued in the previous section, one cannot overlook such a clear reference to the constitutional dimension, simply because the form of internet bills of rights is not reflected in their legal force. These documents adopt a constitutional tone to become part of a specific conversation. Internet bills of rights tend to evoke the normative style of constitutional texts because their message is intrinsically constitutional. By adopting a constitutional tone, internet bills of rights aim to make visible the ongoing process of defining constitutional values of the digital society that would otherwise risk remaining concealed in the complex meanders of corporate policies, in the impenetrability of the code of software, and in recondite conclaves of national and international politics. These documents strive to bring the conversation on digital rights and

---


113. Cf. Teubner, *supra* note 37, (author highlights the latent character of the process of emergence of civil constitutions in comparison with the “blinding glare” of political constitutions. Reflecting on Teubner’s point, we could argue that the Internet bills of rights, by evoking the dimension of political constitutions, would indirectly reflect part of this glare, becoming more visible than the concealed processes of definition of constitutional norms of private corporations, for example).
principles to the attention of global citizens. The language of constitutions is traditionally closer to civil society than the obscure jargon of ordinary legislation and corporate terms.\textsuperscript{114} Embracing a constitutional tone is therefore a way to enhance the communicability of a message.\textsuperscript{115}

At the same time, internet bills of rights adopt the \textit{lingua franca} of constitutional rights and principles to nourish a debate on a specific topic and to address a particular audience. These documents convey through a common language a series of societal instances, ultimately unifying a discourse otherwise fragmented in the multiplicity of normative dialects of corporations and institutions.\textsuperscript{116} The objective of internet bills of rights is to foster a conversation on how to generalise and respecify core constitutional values in the context of the digital society and the language of constitutions offers a fertile field of exchange to this end, allowing the interlocutors to reflect and react by using the same language.\textsuperscript{117}

In a time when the digital revolution generates a climate of legal incertitude and axiological dilemmas in relation to core constitutional aspects, internet bills of rights aim to develop orienting principles.\textsuperscript{118} Traditionally, constitutions are created to transmit foundational societal values from generation-to-generation. Internet bills of rights similarly seek to articulate lasting guidelines that could work as reference points in the future: the most valuable normative contribution in this fast-

\textsuperscript{114.} See generally Giovanni Sartori, \textit{Constitutionalism: A Preliminary Discussion}, 56 AM. POL. SCI. REV. 853 (1962) (discussing the positive meaning that the notion of constitution has in our minds).

\textsuperscript{115.} Cf. \textit{EUROPEAN CONSTITUTIONALISM BEYOND THE STATE}, 2 (J. H. H. Weiler & Marlene Wind eds., 2003) (explaining that the idea of a Constitution for Europe became suddenly popular because the constitutional language was “a fashionable code work, like ‘governance’, for the need to engage in a more profound institutional reform in view of enlargement.”)

\textsuperscript{116.} See Evgeny Morozov, \textit{The Case for Publicly Enforced Online Rights}, FINANCIAL TIMES (2018), https://www.ft.com/content/5e62186c-c1a5-11e8-84cd-9e601db069b8 (last visited May 22, 2019) (arguing that digital rights would not be expressed in the texts of our constitutions, but would rather be the result of a “bundle of permissions of technology platforms.”)


\textsuperscript{118.} See \textit{LESSIG}, supra note 11 (for a discussion of the “latent ambiguity” of constitutional texts); see also Casacuberta, Senges & Duart, \textit{supra} note 75 (for an analysis the problem of rights clashes in the context of the digital society).
changing digital society.\textsuperscript{119}

Lastly, the language of constitutions purveys a rich cultural baggage. Internet bills of rights can benefit from concepts, terminology, ideals, values, and mechanisms that have been developed over centuries to address issues of the protection of individual rights and the balancing of power.\textsuperscript{120} Evoking this conceptual universe means that the conversation of internet bills of rights posits itself in a precise scenario instead of starting from scratch. It aims to continue an old discourse relating to rights and powers, and this time involving more people, addressing a more composite audience and speaking to global actors.

3. \textit{Gradualism}

Internet bills of rights, of course, are not printed texts or manuscripts signed with ink and sealed with wax. These modern declarations of rights are simply published online, a practice so easy and accessible today, that really everyone, from civil society groups to single individuals, can draft one’s own charter and spread it on the web. It is, therefore, not surprising to see a broad array of non-institutional actors involved in initiatives seeking to draft these documents. Nor is it unexpected to observe a high number of internet bills of rights. The "parchment" of these declarations is largely available and affordable, and this partially explains why so many internet bills of rights have emerged over the past few years.

Part of the scholarship has criticized such a "fragmented" and "uncoordinated" way in which the discourse on digital constitutionalism is conveyed by Internet bills of rights.\textsuperscript{121} However, in light of the experimentalist value of these documents that has been emphasised in the previous sections, considering their purpose of nourishing a discussion on rights and principles for the digital age through an accessible and normative language, it is possible to read the multiplicity

\textsuperscript{119} See RODOTÀ, \textit{supra} note 60, at 60 (discussing the necessity to establish orienting principles instead of detailed norms that could not follow the pace of societal changes in the digital society); cf. Martti Koskenniemi, \textit{Constitutionalism as Mindset: Reflections on Kantian Themes About International Law and Globalization}, 8 THEOR. INQ. L. 9, 13 (2006) (for a discussion of "deformalization" as the process whereby the law retreats solely to the provision of procedures or broadly formulated directives to experts and decision-makers for the purpose of administering international problems by means of functionally effective solutions and "balancing interests.")

\textsuperscript{120} See TEUBNER, \textit{supra} note 37 (for a recognition of the significance of political constitutions as "historical model for civil constitutions." Teubner however warns against the risks of "over-hastily" transposing the "stock of historical experience" of political constitutions, overlooking the peculiarities of the global society.)

\textsuperscript{121} Yilma, \textit{supra} note 75, at 126.
of internet bills of rights as a positive form of gradualism. These declarations are progressively advancing the discourse on constitutional values for the digital society. They are testing new solutions and contributing to a debate; they are dialoguing and generating dialogue.

Teubner, developing his theory of societal constitutionalism, wrote:

Civil constitutions will not be produced by some sort of big bang, a spectacular revolutionary act of the constituent assembly on the American or French model. Nor do the global regimes of economy, research, health, education, the professions have a single great original text embodied as a codification in a special constitutional document. Instead, civil constitutions are formed in underground evolutionary processes of long duration in which the juridification of social sectors also incrementally develops constitutional norms, although they remain as it were embedded in the whole set of legal norms.\(^\text{122}\)

Internet bills of rights are halfway between political constitutions and Teubner’s civil constitutions. On the one hand, they adopt the language of political constitutions and evoke the constitutional dimension to become a visible discussion on how to translate constitutional principles in the digital society. On the other hand, internet bills of rights emerge outside institutionalised political processes, as they are issued from novel combinations of institutional actors, civil society, corporations, and single individuals. In a similar way to Teubner’s civil constitutions, internet bills of rights do not boastfully claim their primacy as if they were the unique charter of rights for the internet. There is no single, solemn constituent assembly in charge of drafting the internet constitution. Internet bills of rights emerge with the awareness to add, each one singularly, a small brick to the conversation on digital rights and principles. There is no doubt that the level of visibility, be it economic or social, of the promoters of these initiatives may affect their specific weight in such a debate. Yet, one could argue that, if comprehensively regarded, every internet bill of rights increments and nourishes this dialogue, without pretending to declare the final word.\(^\text{123}\)

\(^{122}\) Teubner, supra note 33.

\(^{123}\) Cf. About the People’s Communication Charter, http://www.pccharter.net/about.html (last visited Oct 6, 2019) (“The People’s Communication Charter is a first step in the development of a permanent movement concerned with the quality of our communication environment. [...] The Charter is not an end in itself. It provides the basis for a permanent
As briefly outlined above, Azzariti advocated for a similar pluralism of initiatives in order to reach a consensus on a constitution for the digital society. He criticised the idea of creating, in the present phase, a legally binding constitutional text articulating digital rights and principles. Such a global lex superior would risk becoming a form of “cultural colonialism” and would tend to weaken the power of existing constitutional instruments. Conversely, the appropriate route to follow is, for Azzariti, a “laborious and slow” process of defining foundational principles for the digital society through the elaboration of non-binding international charters, promoted by a multiplicity of stakeholders and having a persuasive value. In this way, he rightly observes, “building a virtual democracy could come back to be a political objective of real citizens.” Indeed, the multiplicity of internet bills of rights is not only a positive factor in allowing for experimentalism and gradualism in the discourse on digital rights and principles, but also because such pluralism implies a higher level of involvement by civil society and, ultimately, greater democracy.

Gradualism, of course, naturally entails time. Internet bills of rights do not generate a coherent and immediately-applicable set of principles, as—borrowing Teubner’s words—in a sort of normative “big bang.” They slowly advance a reflection, offering stakeholders involved the possibility to cogitate about the most appropriate way to translate core values and principles of contemporary constitutionalism in the context of the digital society. Internet bills of rights are not the output of a multiplicity of Solon, the Athenian lawmaker who drafted a new constitution for his city and then voluntarily went into exile to avoid the modification of his own laws. The authors of internet bills of rights propose their own declarations to partake in a dialogue; they do not leave the constitutional agora. The emergence of each of these texts does not singularly represent a transient constitutional moment. Internet bills of rights are part of an unfinished, continuous phenomenon of cogitation about digital rights and principles. A Sisyphean labour—one could say—but not in a pessimistic sense. The journey that leads to the definition of the constitutional principles for the digital age is certainly more relevant than the ephemeral instant of their solemn proclamation. In this context too, the lesson taught by Albert Camus in The Myth of Sisyphus is true: “[t]he struggle itself toward the heights is enough to
fill a man’s heart. One must imagine Sisyphus happy.”

The extended temporality of the phenomenon of emergence of internet bills of rights not only allows for gradual experimentalism and enhances the level of cogitation, but it is also a necessary ingredient to generate far-sighted constitutional principles. The declarations of internet rights aim to translate the core values and ideals of contemporary constitutionalism in a way that they can speak and last in the context of a fast-changing digital society. It would be of scarce utility to elaborate a series of principles exclusively tailored to the current societal scenario. The speed of digital years no longer seems to reflect the Gregorian Calendar’s units, but it is somehow similar to a cat’s age. In one calendar year, the digital society moves forward six times. However, we do not have any crystal ball to foresee how our society will appear in the coming decades. It is then crucial to craft future-proof constitutional rules. They do not have to last forever. It would be naive to think it possible to do as Solon did in Athens: to establish our law and go into exile to never modify it. However, taking time to understand what the core principles of contemporary constitutionalism are and how to re-specify them in the context of the digital society is a good strategy to elaborate a farsighted and future-proof set of norms.

Espousing this perspective, the gradual process of emergence of internet bills of rights is not necessarily disoriented or disorganised. It is running a long “slow and steady” race, in the hope of achieving a better and long-lasting result. If history can teach us something, it is worth remembering Cicero’s words:

Cato used to say that our constitution was superior to those of other States on account of the fact that almost every one of these other commonwealths has been established by one man, the author of their laws and institutions; for example, Minos in Crete, Lycurgus in Sparta, [...] ; on the other hand, our own commonwealth was based upon the genius, not of one man, but of many; it was founded, not in one generation, but in a long period of several centuries and many ages of men.129

129. MARCUS TULLIUS CICERO, DE RE PUBLICA (Clinton Walker Keys tran., 1928), Book II, I, 2; 110-113.
By highlighting the peculiar features in which the force of internet bills of rights lies, we gain a better understanding of the role these declarations play within the constitutional ecosystem. Internet bills of rights do not emerge in a legal vacuum, but are an integral part of a conglomerate of constitutional mechanisms that are seeking to translate our constitutional values and principles in the context of the digital society. In contrast to most other constitutional instruments, often caged in rigid institutional processes or subject to the influence of a single societal actor, these declarations represent a ductile instrument allowing a multiplicity of stakeholders to participate in a reasoned conversation on digital rights and principles. Internet bills of rights adopt the lingua franca of constitutions to foster a debate, gradually experimenting new ideas through a cogitated process of generalisation and re-specification.

Anne Peters regarded the emergence of transnational constitutional instruments as an expression of “compensatory constitutionalisation on the international plane.”130 Similarly, in light of what has been argued in the previous sections, one could posit that internet bills of rights emerge to play a “compensatory” role within the constitutional ecosystem.131 Freed from institutional constraints, they aim to complement the lack of plasticity, the slow and rigid institutional processes of deliberation, the innate caution and absence of experimentalist tendency of other constitutional mechanisms by courageously promoting the conversation on digital rights and principles. Internet bills of rights surface to counterbalance a static and conservatory tendency within the constitutional ecosystem. They represent—one could say—the most innovative layer of the process of constitutionalisation of the digital society.

However, compensation should not be intended as implying that internet bills of rights and other instruments composing the constitutional conglomerate are de facto airtight containers. As previously seen, the constitutional ecosystem is more similar to a porous material which allows its internal fluids to circulate as in a system of

communicating vessels. The relationship between internet bills of rights and other constitutional instruments is therefore not only compensatory, but—one could add—“stimulatory” at the same time. The tissue of the other constitutional mechanisms, such as national and supranational constitutions, judicial decisions and laws or internal rules of private actors, is permeable and, although partially soaked, it reacts when it comes in contact with the innovative fluids synthetized in the internet bills of rights’ conversation.

These declarations are pushing forward the debate on how to shape the digital constitution by translating the core principles of contemporary constitutionalism in the context of the digital society. The ideas that emerge from this exchange unavoidably come in contact and, potentially, persuade other societal layers, such as politics or the economy, that the time has come and sufficient consensus has been reached to incorporate them in their own constitutional mechanisms. Internet bills of rights can therefore play a litmus test function within the constitutional ecosystem. When compared to other constitutional instruments, these declarations allow for the measurement of its health. They highlight areas of constitutional “anaemia,” in which the discrepancy between legal rules and social reality has reached its apex, and consequently stimulate the emergence of constitutional counteractions in other parts of the constitutional ecosystem.

132. See Graber, supra note 109, at 551, (for an argument that the formal and social constitution should be regarded as two “interconnected vessels”); see also Celeste, supra note 83; cf. Teubner’s conception of porous law in GUNTHER TEUBNER, LAW AS AN AUTOPOIETIC SYSTEM (1993).

133. See CELESTE, supra note 1.

134. See id. at 13.
The Normative Power of Artificial Intelligence

GIOVANNI DE GREGORIO*

ABSTRACT

Artificial intelligence technologies are spreading across society. Generative systems, such as ChatGPT and DALL-E, provide only some examples of the expanding consumption and commodification of artificial intelligence applications in daily life. Nonetheless, the extensive trust and reliance on these technologies in public and private sectors is raising questions for the rule of law. Artificial intelligence technologies are not only mere tools which challenge the protection of fundamental rights when these systems moderate online speech, check employment performances in the workplace, and evaluate credit scores. Particularly, machine learning technologies also contribute to creating norms and rules shaping the enforcement of their functions, thus defining another generative layer of normativity competing with the rule of law in the algorithmic society. This work argues that artificial intelligence systems, particularly machine learning, develop norms by experience and learning within an opaque, technical space. The norms governing these systems are not always immutable but shaped across time. In the algorithmic society, code is not only law but also a source of law. The consolidation of this normative power, or the rule of tech, raises questions for constitutional democracies that are already struggling with solutions to limit other forms of normativity, primarily the predominance of online platforms in the setting of transnational private standards. This plurality of sources has put the rule of law under pressure. The expansion of the rule of tech as a source of norms leads to addressing the spaces for the rule of law and the limits of powers in the algorithmic society, as underlined by the European regulatory approach on artificial intelligence. Within this framework, this work analyses the challenges raised by the normative power of artificial intelligence systems and examines the spaces for the rule of law in the algorithmic society.

* PLMJ Chair in Law and Technology, Católica Global School of Law, Universidade Católica Portuguesa, Lisbon.
Keywords: digital constitutionalism, artificial intelligence, rule of law, powers, norms

I. INTRODUCTION

The consolidation of the algorithmic society is ongoing, and this process is no longer surprising. Increasingly, public and private actors implement digital technologies that contribute to mediating daily life and society at large, and these interventions have now moved to an online dimension as the overlap between the online and offline dimension increases. What is particularly concerning is the general reliance, or delegation, that is driving this shift. Decisions on how to moderate online speech, assign credits, or assess the repetition of a certain crime are increasingly left to the determinations coming from a set of codes and statistical models. And the provision of artificial intelligence technologies without a specific purpose, or General-Purpose Artificial Intelligence systems, which generate text (e.g., ChatGPT) or images (e.g., DALL-E 2), makes this process increasingly visible and within everyone’s reach.

This process has already produced constitutional effects. In the United Kingdom, the determination of school exams and marks by machine learning has triggered public protest about education. The Dutch tax authority used an algorithm to spot childcare fraud, and this use led to discriminating against certain “non-Western” groups. In the United States, some cases have already underlined forms of discrimination, like in the criminal justice system because of the automated calculation of recidivism rates or in labor practices because

---

of worker monitoring. These examples underline how delegating decision-making to artificial intelligence systems raises constitutional questions not only about the impact of these systems on fundamental rights, and, more broadly, on democratic values, but also on the potential marginalization and depowering of the principle of the rule of law in the digital age. These systems are increasingly called to make decisions that, de facto, are based on (technological) standards embedded in their design, not necessarily aligned with legal standards or the protection of public interest. Indeed, the threats of "algocracy" do not only question the role of humans, or the protection of fundamental rights, but also the role of the rule of law.

Nonetheless, these concerns do not fully exhaust the constitutional questions raised by artificial intelligence systems. Algorithmic technologies are not only instruments to exercise powers, which can interfere with fundamental rights, but can also be considered as rule-makers. Rather than mere executing tools based on pre-settled instructions and standards, machine learning, and deep learning, systems learn how to perform their task and adapt it through experience. In this case, these systems exercise normative powers. For example, the removal of online content is not only the result of community standards of social media or the design of algorithmic technologies but also the ability of artificial intelligence to define and, then, decide what to consider as hate speech, and, only after that, remove that content. Similar considerations extend to the public sector. Digital surveillance underlines how machines compute certain degrees of risk and then flag some cases to law enforcement. Decisions are based not only on data stored by law enforcement authorities or technical standards but also on a set of self-generated technical norms that assess risks changing and adjusting over time.

Whether hate speech content is considered harmful and then removed, or a citizen is profiled as a potential suspect up to a certain

---

percentage of risk, is a decision shaped by a mix of technical norms and machines’ discretion which tends to change through experience. The more machines moderate content, the more this process will be shaped by their assessment. Therefore, artificial intelligence systems do not only enforce rules or standards but also adopt and shape norms. The learning process of these technologies could change how the right to freedom of expression prevails over another conflicting right or legitimate interest, such as dignity or public health, even extending or restricting legal notions or private standards of speech. Likewise, generative models provide answers that change over time. Large language models are trained with new information coming from developers and users, which shape the norms defining the output.

This normative power illustrates how saying that code plays the role of the law, or “code is law,”14 does not entirely represent the evolution of the algorithmic society. Code is also a “source of law.” Artificial intelligence systems are not only examples of how code is law but also how code creates law, or rather produces norms. Particularly applications based on machine learning and deep learning provide another generative normative source that shapes the algorithmic society. While public actors regulate digital technologies15 or tech giants16 express their powers by enforcing rights and freedom on a global scale based on their terms of services,17 technology also expresses a form of automated governance far from the logic of public and private actors. This understanding, which can be called the rule of tech, challenges the traditional boundaries of the rule of law, or even the rule of the platform. Therefore, the normativity of artificial intelligence tends to escape public and private oversight, thus raising questions beyond the boundaries of liberal, global, and societal constitutionalism.18

However, political and regulatory reactions to the marginalization of the rule of law are still the exception. The European Union approach provides a paradigmatic example of the attempt to reposition the rule of law at the core of the algorithmic society. The top-down, risk-based approach introduced by the Artificial Intelligence Act tends to limit the

15. See Return of the State?, in POWER AND AUTHORITY IN INTERNET GOVERNANCE: RETURN OF THE STATE? (Blayne Haggart, Jan Aart Scholte, Natasha Tusikov (Eds.) 2021).
18. GUNTHER TEUBNER, CONSTITUTIONAL FRAGMENTS: SOCIOECONOMIC CONSTITUTIONALISM AND GLOBALIZATION 1, 2 (2012).
scope of the rule of tech, and not only in relation to the normative power of online platforms, or, more generally, the private sector. In order to protect European values, the Artificial Intelligence Act aims to provide a competing normative framework. Even if this regulation does not always match the evolution of artificial intelligence technologies as underlined in the case of generative systems, nonetheless, the European policy on artificial intelligence provides an opportunity to consider the spaces for the rule of law in the algorithmic society, thus underlining the constitutional relevance of these challenges.

Within this framework, this work aims to examine the normative power of artificial intelligence. This paper argues that artificial intelligence technologies, particularly in the cases of machine learning and deep learning, contribute to the production and adaptation of the rules characterizing the algorithmic society. The expansion of the rule of tech raises questions for constitutional democracies. As a normative system, the rule of tech competes with other forms of normativity, primarily the rule of law and the rule of platforms. The primary challenge is how to reconcile the normative clash between the rule of law and the rule of tech and how to define the spaces for the rule of law to limit the exercise of unaccountable powers in the digital age.

The intrinsic normative power of artificial intelligence technologies is a call for digital constitutionalism, particularly about the protection of rights and democratic values in the algorithmic society. By looking at the perspectives brought by liberal, global, and societal constitutionalism, this work analyses the normative power of artificial intelligence and examines the spaces for the rule of law in the algorithmic society. This work aims to complement the studies on algorithmic regulation and the rule of law, by looking at how artificial intelligence technologies do not only define a regulatory system based on code and data but also generate another layer of normativity that


22. See, e.g., ALGORITHMIC REGULATION (Yeung Karen & Martin Lodge eds., 2019); Mireille Hildebrandt, Algorithmic Regulation and the Rule of Law, PHIL. TRANS. R. SOC. A. (2018); Roger Brownsword, Technological Management and The Rule of Law, 8 LAW INNOVATION TECH. 100 (2016).
extends beyond the state as well as beyond private oversight.

The first part of this work examines how the shift from law and territory to norms and spaces in the digital age has challenged the predominance of the rule of law. The second part analyzes the normativity of artificial intelligence technologies which contribute to creating and adjusting the norms shaping their activities. The third part addresses the clash between different forms of normative powers and the competition between the rule of law and the rule of tech. The fourth part examines the spaces for the rule of law to address the challenges raised by the rule of tech in the algorithmic society.

II. FROM LAW AND TERRITORY TO NORMS AND SPACES

The consolidation of the digital age has contributed to expanding a plurality of normative sources. Online platforms and standard-setting organizations are only two examples of how the rules governing the digital environment do not exclusively originate from states’ boundaries. Not so different from other global trends, the development of digital technologies has raised questions about the role and scope of the law by defining different patterns of convergence, usually named “globalization,” where the state-centric model has started to lose its power. Even if sovereign claims are still relevant but under pressure, territorial borders are challenged by “a world in which jurisdictional borders collapse, and in which goods, services, people, and information ‘flow across seamless national borders.’” The rise of “global law” defines a meta-legal system where different organizations and entities produce and shape norms with extraterritorial implications. Therefore, norms are not only the result of states’ production but also come from multiple sources on a global scale.

A striking example of the relationship between law and space is found in the digital environment. At the end of the last century,

---

Johnson and Post wrote that “[c]yberspace radically undermines the relationship between legally significant (online) phenomena and physical location.” This statement represents how the gap between law and space and is one of the reasons that critics firmly deny the idea of cyberspace as a new “world” outside the influence of sovereign states. Territorial boundaries are known for their ability to define limited areas where states can exercise their sovereignty. Inside a certain territory, people are expected to comply with the applicable law in that area. Since cyberspace is not a “lawless place,” states can impose their sovereignty, especially by regulating network architecture. Despite the relevance of these positions, while states can exercise their sovereign powers over the digital environment within their territories, other systems contribute to producing their norms in turn. More specifically, this trend in the digital age derives from the code’s architecture which sets the rules constituting the meta-legal norms of the digital environment.

The multiplication of private and technological standards increasingly led to replacing the relationship between law and territory with the interrelation between norms and spaces. Digital technologies have contributed to expanding the sources of norms production that defines the normative order of the internet. Constitutional democracies have not been spared in this shift. Standards to protect freedom of expression or the enforcement of public policies in the digital environment are increasingly left to the discretion of private actors and technological standards on a transnational scale.

Instead, constitutions traditionally embody the values and principles that a specific community adheres to and respects. They represent an expression of the social contract between public power and citizens. Constitutions have seen the light in different contexts and local dynamics through different forms of constituent powers, thus defining an intimate relationship between constitutions and a certain area of space (i.e., territory) over which the sovereign power is exercised and limited.

The relationship between (constitutional) law and space is intricate. The law stands on a territorial space and relies on political processes

legitimizing its creation. Formally, outside the domestic legal framework, other legitimized binding forces over a certain territory do not exist unless tolerated, accepted, or even authorized, by the legal framework itself. Substantially, the law is only one of the systems influencing space. From the perspective of societal constitutionalism, the law is one of the systems interacting with other functional social systems. Although social systems tend to be normatively closed since they autonomously develop their own systems, these systems are cognitively open.\textsuperscript{35} Therefore, law, economics, technology, science, and politics develop their own rules through their institutions, yet at the same time, they can observe their social environment and other systems and be indirectly affected by them. If one focuses on this process of autopoiesis, state-based law is not the only legitimated normative infrastructure in a certain territory anymore. Instead, it is only one of the fragments composing a constitutional puzzle on a global scale. Even in the digital age, law, technology, and society, as examples of social systems,\textsuperscript{36} produce internal norms while continuously shaping each other in a process of mutual influence,\textsuperscript{37} or rather digital constitutively.

The law is the result of its logic as well as the compromise between the technological architecture, social norms, and market forces.\textsuperscript{38} As a result, the law indirectly influences the other systems which, even if characterized by reflexivity, inevitably are part of a greater understanding. Usually, legal categories, such as rules, authority, or rights and freedoms, contribute to shaping the boundaries of recognized powers. These definitions do not exist outside the legal framework but are created within the rationality of the law. However, these systems are not proof of systemic interferences. Likewise, legal systems based on definitions, scope, and enforcement shape the boundaries and characteristics of technology and society.\textsuperscript{39} In other words, the peculiarity of the law as a social subsystem is to define spaces as delegated and autonomous manifestations of powers.

This form of pluralism leads to considering constitutionalism under a broader umbrella where the link between law and territory is increasingly replaced by the relationship between norms and powers coming from different autonomous rationalities that mutually shape each other. By moving from a unitary view of the law, or positivism, to

\textsuperscript{35} Gunther Teubner, \textit{Autopoiesis in Law and Society: A Rejoinder to Blankenburg}, 18 \textit{L.


\textsuperscript{38} Lessig, supra note Error! Bookmark not defined., at 123.

legal pluralism, it cannot be neglected how other systems develop their norms and principles, irrespective of whether they are considered as “law.” Therefore, the relationship between law and territory characterizing states’ boundaries and sovereignty tends to lose its exclusiveness, thus, leaving space for the consolidation of another dyadic relationship between norms and spaces.

The normative power of artificial intelligence adds another generative layer of norms that provide alternative standards to the protection of constitutional values, including the rule of law. This challenge is primarily connected with the consolidation of the rule of tech that complement the idea of “code is law” with the perspective of “code is source of law.”

III. THE RULE OF TECH

The normative power of artificial intelligence technologies is primarily connected to the ability of making decisions and learning from experience. Artificial intelligence systems are no longer relegated to research labs, but they are now spreading across society, as underlined by generative models. The development of these systems would not raise constitutional concerns if they were not left to make decisions on fundamental rights and democratic values based on an opaque self-generative normative power.

At first glance, algorithmic systems appear as neutral technology that can extract values from information and that are useful for society, specifically as public and private actors use them to provide public services or run a business. The autonomy of these machines is also the reason for their role in finding new information and driving scientific discovery. However, algorithms are far from being mere neutral technologies or mathematical models that provide outcomes in a certain form based on the processing of information. Rather, these technologies transform inputs into outputs, and thus, de facto, they express a value judgement. In other words, automated decision-making systems are value-laden. These systems are far from perfect, and they lead to potential discriminatory bias or to the exposure of objectionable content, as demonstrated in the case of content moderation or search

Even large generative models raise this issue considering that their inputs do not merely come from the user’s question but also from the training of data from different sources that could reflect other biases. These technologies make decisions that affect individual and collective rights as well as expectations and trust.

Besides, potential biases are not only the result of data processing but also are embedded by humans who are involved in the phase of programming and developing. As such, human biases and values are reflected in the technologies and their design. Whether an algorithm aims to protect the right to freedom of expression is reflected in how many posts it will keep online. Likewise, an automated decision-making system could be more inclined to profile criminals based on the need to protect public interests over privacy. In some cases such as rule-based systems, these constitutional conflicts are defined in the architecture of these technologies. Therefore, in this case, humans play a critical role as the constitutional creators of algorithmic checks and balances.

Nonetheless, these (algorithmic) decisions are not exclusively based on design choices. Machine learning systems also learn how to perform their tasks, which shape their own activities, or, from a constitutional perspective, the weight of the original system of checks and balances. Such machine determinations mediate rights and freedoms based on a mix of pre-settled and self-generating standards. By processing vast amounts of information and data, artificial intelligence systems do not only lead to complementing, or replacing, legal norms with computing standards, but also create new norms which tend to differ from private standards or states’ regulation.

Within this framework, artificial intelligence systems develop norms leading to a process that is de facto constitutional. In this sense, the technical norms resulting from artificial intelligence applications become constitutionally normative. For instance, when considering social media, algorithms determine what information is displayed first and what is hidden, for instance, recommending a specific journal article or blog post to read. They moderate and curate content based on

46. See generally MIREILLE HILDEBRANDT, SMART TECHNOLOGIES AND THE END(S) OF LAW: NOVEL ENTANGLEMENTS OF LAW AND TECHNOLOGY (2016).
categories, such as hate speech or disinformation. These activities are not fixed in a timeframe but change across time through repetitions and adjustments. In turn, such autonomy leads artificial intelligence technologies to enforce and balance rights and freedoms as well as to shape and develop a system of norms by experience. In a sense, this context is similar to how common law systems create norms through the accumulation of courts’ case law or to how civil law systems adjust their norms to address new challenges.

Unlike the law of nature, in the realm of artificial intelligence it is not always possible to predict the consequences and the forces that shape the creation of these norms. The norms of artificial intelligence are not hidden into an opaque policy framework or the law of nature, such as gravity. There is no democratic oversight on their creation, and these norms also do not answer to the logic of “conditional rationality”—if A happens, then B is expected. Rather, artificial intelligence technologies reflect the condition that if A happens, then B could be expected. Therefore, the reasons for this opacity are not only related to legal barriers or opaque policy-making but also to the rule of math and statistics.

It is precisely here that artificial intelligence systems challenge the role of the rule of law. These technologies do not aim to provide a certain answer but to secure a reasonable outcome. The implementation of these technologies has moved the focus from causality to probabilities and correlations. The limit of traditional systems of processing to deal with the vast amount of data encourages the implementation of statistical methods. This shift from causality to probability is not neutral, and it raises concerns about the reliance on the outcomes of these technologies, particularly when they are used to make or support decisions on fundamental rights and democratic values, as well as about the predictability of their norm creation.

The opacity of artificial intelligence escapes the logic of the rule of law and moves to the realm of statistics and probability based on the large amount of data that makes any expectation of a certain outcome unpredictable. The law is usually based on definitions that aim to ensure uniformity and equality. Instead, algorithmic technologies make decisions without relying on a legal basis. Their decisions are not based on a threshold of illegality but on statistical approaches that lead, for instance, to pattern recognition, clustering, and classification of objects. All these activities are shaped by multiple influences, such as the code,

47. Tarleton Gillespie, Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media 5 (2018) (articulating the idea that platforms have to moderate).
the data used in a certain case, and the learning capabilities of algorithmic technologies that can also lead to the creation of new norms. The definition, enforcement, and balancing of these rules are not only mediated by design choices but also by the power of math, which is the real governor of the rule of tech.

As autonomous agents, these technologies do not align their behaviors based on predictable norms. The balancing of conflicting interests is based on a probabilistic approach that leads to clustering a certain object or defining patterns. The opacity of decision-making does not always reveal which values have guided a certain outcome, thus limiting the possibility of detecting reflexivity. As the law tends to reflect its dynamics, such as injustice or inequality, it is possible to argue that the same is true for algorithms that reflect their technical and social underpinning values. Therefore, in this sense, even if artificial intelligence systems can be considered potentially reflexive, they tend to be lawless.

This consideration also explains why the norms created by artificial intelligence technologies cannot be considered as social norms. The logic of artificial intelligence is not based on the mere acceptance and sharing of rules agreed upon by a community. Although operational standards are defined by programmers and developers while also shaped by service providers and users, there is not a unique community that defines the common standards that govern algorithmic systems. Therefore, rather than social norms, algorithmic technologies create self-generating technical norms with societal constitutional implications. This context seems particularly relevant when considering the implementation of increasingly autonomous technologies in the algorithmic sphere.

Within this framework, the normative power of artificial intelligence leads to examining the remaining spaces for the rule of law in the algorithmic society. The rule of tech is another example that underlines how the rule of law is under pressure in the digital age and how its realm is limited by the expansion of different spaces that produce norms. The norms shaped by state or private normativity do not exhaust the forces which overlap and compete in a process of mutual influence. Therefore, the position of the rule of law can be examined by looking at the relationship between different layers of normativity in the digital age.

49. Jennifer Cobbe, Legal Singularity and the Reflexivity of Law, in IS LAW COMPUTABLE?: CRITICAL PERSPECTIVES ON LAW AND ARTIFICIAL INTELLIGENCE, 1, 16 (Simon Deakin & Christopher Markou eds., 2020).

50. See generally ERIC POSNER, LAW AND SOCIAL NORMS (2000).
IV. COMPETING NORMATIVE POWERS IN THE ALGORITHMIC SOCIETY

The scope and effectiveness of the law in the digital environment has been questioned since the advent of the internet. Constitutional democracies have highly struggled with extending their powers, and laws, on a global scale beyond their territorial boundaries as particularly underlined by the role of courts. These challenges have led to relying on alternative systems to enforce public policies online, particularly by delegating private actors to serve such a role.

The threats to the principle of the rule of law are linked not only to the abuse of public powers or the consolidation of the rule of tech but also to the possibility that private actors have developed a set of private standards that clash with public values, precisely when their economic freedoms turn into forms of power. This shift has raised questions about the rule of law that meets some limits when walking outside the public realm and moving to the private sector. The rule of law is a guide for public actors to ensure equal treatment before the law, and it protects rights and freedoms and limits the abuse of powers by unaccountable authorities holding decision-making powers. The principle of the rule of law constitutes a guide for public actors when interfering with rights and freedoms. This understanding also applies in the field of digital technologies that public actors can use as instruments of social control.

In other words, the space of the rule of law is not only limited by the consolidation of the rule of tech but also by the rise of private normative powers that are usually expressed through self-regulation.

This trend can be analyzed from multiple perspectives in the digital age. In the case of social media, these actors set and enforce their standards of protection while balancing conflicting interests at stake, thus acting as governors. The deplatforming of former President Donald Trump in the aftermath of his support of the violent attack on the Capitol is only one of the examples that underline how platforms

have consolidated their role as gatekeepers of information.\textsuperscript{57} The governance of this form of decision-making is not shared but rather centralized and covered by unaccountable purposes. In this case, rights are established by a private form of authority through opaque terms of service contracts that are designed and enforced by private actors.\textsuperscript{58} These spaces are then shaped by private actors—in this case, platforms—who impose their private normative powers over the rule of law.

The competition between public and private normative powers does not exhaust the plurality of sources meeting and overlapping in the digital age. The rule of tech defines another layer of normativity. Checking whether automated decision-making technologies are aligned with legal norms or the values and principles defined in their technological design is not always possible. The lack of transparency and accountability of these systems challenges the possibility of monitoring for public actors as well as for private actors implementing these technologies. The problem of black box and biases are only two of the most important examples of the pitfalls relating to artificial intelligence,\textsuperscript{59} as underlined by the increasing centrality of explanation and interpretation.\textsuperscript{60} Machines are still not entirely able to interpret real dynamics and exactly understand contexts and emotions,\textsuperscript{61} and it is likely that they will need some form of human support.

This form of technological regulation is different from legal regulation. Technological regulation is not the result of a democratic process; it excludes disobedience, and it does not allow contest due to the lack of transparency and accountability of decision-making. These decisions escape even the scrutiny of their developers and programmers, and the scale of artificial intelligence applications does not make oversight profitable, or even possible, to check such norms production through a traditional form of (judicial) review. This risk would lead to limiting the application of artificial intelligence in areas of public interest such as justice and legal adjudication, but it also raises


\textsuperscript{59} Tal Zarsky, Transparent Predictions, 4 U. ILL. L. REV. 1503, 1507 (2013).

\textsuperscript{60} John Zerilli, Explaining Machine Learning Decisions, 89(1) PHILOSOPHY OF SCIENCE 1 (2022).

\textsuperscript{61} Andrew McStay & Lachlan Urquhart, This Time with Feeling? Assessing EU Data Governance Implications for Out of Home Emotional AI, 24 FIRST MONDAY 4–5 (2019).
questions about the rule of law in the algorithmic society.⁶² Even technological solutions to solve these challenges do not provide an exhaustive approach. For instance, the decision-making process of artificial intelligence can be subject to an adversarial mediation.⁶³ Rather than designing artificial intelligence systems able to understand the rule of law or the protection of fundamental rights, the goal is to introduce other digital agents that can counterbalance conflicting interests in automated decision-making. This trend toward the use of adversarial systems, which is particularly relevant in the field of machine-to-machine communication, also reflects the tendency of constitutional democracies to recognize a central role in the exchange of different positions in cases involving conflicting interests.⁶⁴ In this sense, as courts ensure judicial review and adjudicate cases based on adversarial systems, artificial intelligence would meet another automated voice that could balance other values in the decision-making process. In a sense, adversarial systems could lead to a process of constitutionalizing artificial intelligence by introducing checks and balances.

The primary concern related to the implementation of adversarial systems is related to the use of technology to solve technology. If it is still not possible to explain the reasons for algorithmic outcomes in a certain case, the primary question is whether it is possible to rely on another opaque system to make an outcome fairer. This problem is not only linked to the technical limits of computing the law but also to the ability of artificial intelligence systems to create norms that inevitably would shape the weights of checks and balances that aim to ensure a fairer process. One solution would be based on a public system of adversarial artificial intelligence.⁶⁶ This approach would rely on algorithmic technologies that are developed or overseen by public actors to check and proceduralize automated decision-making processes, also supporting the collection of data about the adversarial outcomes and the adjustment of algorithmic systems to the rule of law. However, adversarial systems can also amplify biases, and the normative power of these technologies could make adversarial artificial intelligence

---


ineffective due to the changing set of norms developed through experience.

The limit of adversarial systems, or other forms of oversight and review, also results from the predominance of private actors, and their normativity, in developing and implementing automated decision-making systems, particularly when there is no regulation requiring private actors to pursue public interests and respect fundamental rights. The rule of tech does not result from a democratic and dialectic process but from an opaque mix governed by technical determinations. Individuals are increasingly surrounded by technical systems that influence their decisions without them having the possibility to understand or control this phenomenon. The mediation of automated technologies increasingly leads users to participate in a “modulated democracy.”67 Democratic values in the digital age are likely to be mediated by the implementation of artificial intelligence systems as well as by the production of private and technical norms outside any form of public scrutiny.

The primary question is about understanding the spaces for the rule of law in this plural and networked expression of normative powers. The power to generate norms outside any oversight raises concerns for constitutional democracies. This question is central for mitigating the risk that constitutional values, primarily fundamental rights and democratic values, are driven by unaccountable logic resulting from the power of the private sector and the lawlessness of artificial intelligence systems.

V. The Spaces for the Rule of Law

The consolidation of the rule of tech has led to the contraction of the spaces for the rule of law. This situation has increased the pressure on constitutional democracies to address the marginalization of the rule of law in the algorithmic society. Accountability, transparency, and explainability have been characterizing the debates on the law of artificial intelligence.68 Thus, these principles have become the most

important parts of algorithmic regulation.  

This attention shows how constitutional democracies are at a crossroads. On the one hand, public actors can leave artificial intelligence technologies to exercise their normative power, or, on the other hand, they can limit, or even ban, these systems, as underlined by the questions raised by biometric technologies in public spaces. Both these cases lead to constitutional tensions, particularly in terms of effectiveness and proportionality. In the first scenario, constitutional tolerance could amplify the normative power of artificial intelligence outside any public or private oversight. In the second scenario, a general ban could not only be ineffective due to potential technological circumvention but also interfere with other constitutional interests, primarily economic freedoms, thus slowing down the development of algorithmic technologies and increasing the uncertainty about their implementation. The ban of ChatGPT by the Italian Data Protection authority has underlined the limit of generalized bans on artificial intelligence technologies.

Addressing this constitutional challenge leads constitutional democracies to decide whether to accept or react to a plurality of normative powers that compete with the rule of law. The algorithmic society requires constitutional democracies to strike a balance between innovation and the protection of rights in the algorithmic society. The question is not merely how to regulate artificial intelligence but how to accommodate conflicting constitutional poles based on the rule of law as the primary guidance.

The normative power of artificial intelligence can be mitigated by expanding the rule of law in the digital age. The tolerance of constitutional democracies tends to spontaneously trigger plurality, and, therefore, the rise of new normative powers that compress the spaces of the rule of law. When this compression leads to constitutional challenges, it is for the rule of law to expand its role and mitigate the expansion of unaccountable normative powers. This approach leads to adopting a legal framework in the algorithmic society that, first, requires public and private actors to follow procedural safeguards when delegating the setting of standards and decision-making to artificial

---

intelligence systems, and, second, to provide remedies against the effects of automated decision-making on fundamental rights and democratic values. Rather than merely leaving artificial intelligence systems free to make decisions or banning potential applications, this approach focuses on designing *ex-ante* safeguards and *ex-post* remedies that mitigate the challenges raised by the rule of tech.

The process of delegating other agents the responsibility to set standards or make decisions does not define a new trend for constitutional democracies. Powers have been delegated across multiple entities throughout history, and decision-making has often been transferred from the public realm to the private sector.\(^72\) This process is still relevant in the digital age, as underlined by the neoliberal approaches adopted by constitutional democracies that have rendered online platforms private makers of standards and enforcers on a global scale.\(^73\) Boyle already wondered whether the internet would have led to a transformation that challenges basic assumptions not only concerning economics but also constitutional and administrative law.\(^74\) The rise of digital private powers can primarily be considered the result of an indirect delegation of public functions. The shift from public to private in the digital environment is not an isolated case, but rather, it is the result of a general tendency towards the transfer of functions or public tasks from lawmakers to specialized actors in both the public and private sectors.

This trend is part of a larger system of delegation that no longer exclusively involves the relationship between the lawmaker and the government (legislative-executive) but also with two new branches, respectively, public bodies such as agencies (fourth branch)\(^75\) and private entities dealing with delegated public tasks (fifth branch).\(^76\) The delegation of public functions is not merely a unitary phenomenon; it can include agreements between public and private actors based on public-private partnership schemes where private entities provide goods


or services.\textsuperscript{77} The cases of smart cities or governmental services are examples of the shift of responsibilities from the public sector to private entities through instruments of public procurement.\textsuperscript{78} In other cases, the delegation of public functions consists of the creation of new (private or public) entities to perform public tasks, such as the provision of products and services or support to rulemaking activities. In these cases, the establishment of a new government, corporation, or agency is one of the most evident examples.\textsuperscript{79}

The consolidation of the rule of tech amplifies the questions about the limits of delegating powers in the digital age, not only across public and private actors but also to normative technological systems that challenge oversight. In this case, the normativity of artificial intelligence systems tends to overcome the dichotomy of the public/private divide, thus making some of the constitutional instruments ineffective in addressing unaccountable powers. For instance, the horizontal effect of fundamental rights has a limited reach in this case,\textsuperscript{80} particularly considering that it is not possible to require artificial intelligence to respect fundamental rights. Moreover, even if it is possible to apply this approach to private actors who implement these technologies, the normative power of artificial intelligence technologies could limit the possibility to intervene to protect rights and freedoms, or even to ensure constitutional values such as due process and proportionality.

Procedural safeguards in the digital age, such as adequate notice given to individuals affected by the decision-making process, have already been advanced in recent years.\textsuperscript{81} The case of algorithmic impact assessment is only one example of a procedural safeguard that can reduce the consequences of unaccountable delegation. The Council of Europe, particularly the ad hoc committee on artificial intelligence, has focused on the introduction of algorithmic impact assessment to increase the accountability of the public and private sectors when

\textsuperscript{77} See generally Albert Sánchez Graells, Public Procurement and the EU Competition Rules (2d. ed. 2015).
\textsuperscript{79} See generally Marta Simoncini, Administrative Regulation Beyond the Non-Delegation Doctrine: A Study on EU Agencies (2018).
implementing artificial intelligence technologies. Likewise, the possibility for individuals to access remedies, such as judicial review, is another critical step to increase accountability, as underlined by the US Blueprint for an Artificial Intelligence Bill of Rights. The relevance of remedies in the algorithmic society is also connected with the centrality of humans in the algorithmic society.

However, the tolerance of constitutional democracies to other expressions of normative powers deeply influences regulatory approaches. For instance, the liberal perspective driving the rise of online platform powers across the Atlantic has underlined a trend towards polarization in the last twenty years. From the first period of regulatory convergence based on neoliberal positions at the end of the last century, the United States and the European Union have taken different paths. While the EU has slowly complemented its liberal imprinting with a constitutional democratic strategy, as underlined by the adoption of the General Data Protection Regulation, the Digital Services Act, and the Digital Markets Act, the US constitutional framework has not expressed the same concern and instead follows an opposite path. For instance, the Communication Decency Act still immunizes online intermediaries, including modern online platforms, from liability when moderating users’ content. In the field of data, apart

---

82. Ad Hoc Committee on Artificial Intelligence: Possible elements of a legal framework on artificial intelligence, based on the Council of Europe’s standards on human rights, democracy and the rule of law, Council of Europe (Dec. 3 2021), https://rm.coe.int/cahai-2021-09rev-elements/1680a6d90d.
from some national attempts, no harmonized approach exists to privacy and data protection at the federal level. In the fields of both content and data, the US policy is still anchored to a digital liberal approach that considers the First Amendment as the primary beacon of the algorithmic society.

Moreover, constitutional democracy is not only influenced by the challenges raised by legal pluralism in the digital age but also driven by the interest to ensure competitiveness in the technological sector in the long run. Restricting or even banning artificial intelligence systems can slow down the growth and development of digital products and services. In other words, granting extensive protection of fundamental rights over innovation could make constitutional democracies “standard-takers” rather than “standard-makers” in the field of artificial intelligence. Furthermore, if it is true that the market and democracy are intimately connected, the choice of constitutional democracies to limit the normative power of artificial intelligence is driven more by the threats of competitive disadvantages than the protection of fundamental rights and democratic values.

Considering the role of artificial intelligence for the fourth industrial revolution, this is not a trivial constitutional issue. The constitutional advantage of protecting fundamental rights and democratic values in the short term could lead to a situation of de facto technological disadvantage. The need to rely on systems that are developed in areas of the world where the lack of restrictions and liberal approach leaves spaces for the development of unaccountable models of governance would affect the same constitutional rights driving the restrictive approaches to the development of artificial intelligence systems. In other words, the price to expand the rule of law in the algorithmic society could lead to making the rule of tech stronger in the long run.

VI. A REGULATORY ATTEMPT

The European strategy is an example of the struggle with repositioning the rule of law in the algorithmic society. Even before the adoption of the Artificial Intelligence Act, the EU had adopted critical steps. For instance, the Digital Services Act underlines a paradigm shift in the EU by limiting online platforms’ normative powers through procedural safeguards. These measures could be considered not only

91. See e.g., CAL. CIV. CODE § 1798.100 (West 2018).
an attempt to adapt the digital economy to European goals but also a reaction against the consolidation of other forms of normative powers. The Commission has also launched a declaration of rights and principles in the digital age, which defines a human-centric approach guiding the European digital transition. The declaration aims to put people at the center and foster solidarity and inclusion, freedom of choice, participation in the digital public space, safety, security and empowerment, and sustainability. It is an expression of the consolidation of digital constitutionalism, thus underlining how the digital future of the EU will likely be based on a digital compass guided by European (constitutional) values. As stressed by the Commission’s President, Ursula von der Leyen: “We embrace new technologies. But we stand by our values.”

The Artificial Intelligence Act is a critical part of this framework. The objective is not only to promote the development of artificial intelligence technologies in Europe to foster the development of the internal market but also to protect European values. This approach aims to avoid the misuse of technologies that produce risks for public interests and rights that would “contradict Union values of respect for human dignity, freedom, equality, democracy and the rule of law and Union fundamental rights, including the right to non-discrimination, data protection and privacy and the rights of the child.” This duality of goals is precisely the characterization of the European approach at the intersection between digital humanism and digital capitalism.

The Artificial Intelligence Act provides a first, even if not perfect, approach to address the constitutional questions raised by the normative power of artificial intelligence systems. It introduces layers of risks that limit the possibility of implementing certain artificial intelligence systems that are considered “unacceptable” and requires providers of these technologies to introduce procedural safeguards for

---

94. De Gregorio, supra note 18, at 1–2.
97. Artificial Intelligence Act, supra note 17, at 21.
systems that are “high risk” and “low risk.” These layers are defined from the top by the European Commission, thus limiting the flexibility of this legal instrument. The list of “unacceptable,” and therefore prohibited, artificial intelligence systems or even high-risk systems is directly set by the law and is independent of any a posteriori risk assessment by providers or users of those systems.

Even if this approach is rooted in the protection of European values, it could not ensure enough flexibility in the long run, and it also increases legal uncertainty, thus depowering the repositioning of the rule of law in the digital age, as underlined by generative applications such as ChatGPT. The risk-based approach can be complemented by tightening public and private actors through coregulation. The possibility of finding common goals and reach a compromise is another possibility for the rule of law to count in shaping the regulation of artificial intelligence. This approach can also contribute to ensuring that constitutional values are considered in the underpinning rules defining these systems while also ensuring that private actors have margins of discretion in implementing these technologies by following a certain constitutional frame.

Nonetheless, the challenges to the rule of law are not only resulting from the structure of the Artificial Intelligence Act but also from interrelation with other legal instruments, particularly the General Data Protection Regulation (GDPR). For instance, the obligation of data controllers to conduct data protection impact assessment in certain cases tends to overlap with risk management systems for high-risk artificial intelligence technologies. This situation primarily comes from the intimate connection between (personal) data and artificial intelligence technologies. Likewise, the Digital Services Act requires very large online platforms to conduct risk assessments in the process of content moderation, which would also include the role of artificial intelligence technologies implemented to tackle harmful content.

Furthermore, the Artificial Intelligence Act fails to deliver remedies to mitigate the consolidation of the rule of tech regarding the impact of algorithmic technologies on society. This legal instrument does not focus on empowering individuals by providing remedies against the impact of automated decision-making. Even if European values play the role of guiding the interpretation of the Artificial Intelligence Act, they do not fully permeate the top-down classification of risks defined by the Commission based on different artificial intelligence technologies. Unlike the approach followed in the GDPR and the Digital Services Act,

100. Id. at 2.
101. Digital Services Act, supra note 81, at 32.
the Artificial Intelligence Act provides a top-down approach where
public actors define norms that shape the possibility of implementing
artificial intelligence technologies. It does not leave space for redress
against the impact of artificial intelligence systems. Rather than
introducing remedies to protect European values, such as dignity, the
Artificial Intelligence Act looks at these technologies more as products
than as a process of decision-making expressing a normative power that
could affect fundamental rights and democratic values.

The lack of remedies in the Artificial Intelligence Act can also be
understood by considering this legal instrument as a form of protection
of European values against the expansion of technological delegation by
external actors. Rather than providing remedies, the EU aims to
provide legal standards for artificial intelligence that clarifies typologies
of these technologies that are aligned with European values. In other
words, the expansion of the rule of law aims to reduce the consolidation
of the rule of tech beyond Europe. The EU has already shown its ability
to influence global dynamics, also known as the “Brussels effect”, particularly by extending its “regulatory soft power” to other areas of
the world in the field of technology. This approach also comes with a
narrative about digital sovereignty. As underlined by the Commission, “European technological sovereignty starts from ensuring
the integrity and resilience of our data infrastructure, networks and communications” aimed to mitigate “dependency on other parts of the
globe for the most crucial technologies.” This understanding does not
entail closing European boundaries to a form of constitutional
protectionism but ensuring Europe’s ability to define its own rules and
values in the digital age.

The EU strategy is an example of the critical role of the rule of law
in limiting the expansion of the rule of tech. The consolidation of the
normative powers of artificial intelligence leads constitutional
democracies to design safeguards and remedies to address the exercise
of unaccountable powers and limit the impact of these systems on
fundamental rights and democratic values. While the Artificial
Intelligence Act defines a critical step, and reaction, against the
consolidation of the rule of tech, it recognizes that certain decisions
cannot be left to the complete discretion of artificial intelligence and

104. Id.
introduces layers of risks that trigger procedural safeguards. This European approach underlines a critical paradigm shift from neoliberal approaches and techno-deterministic promises to a democratic strategy that aims to expand the role of the rule of law over other forms of normativity.

VII. Conclusions

The consolidation of the algorithmic society is primarily connected to an increasing reliance on artificial intelligence systems. These systems are spreading across society and implemented to make decisions that can also affect fundamental rights and democratic values. Algorithmic technologies increasingly mediate rights and powers, thus challenging the role of the rule of law in the digital age. The normative power of artificial intelligence defines another generative system of norms that meets and competes with other sources, primarily the rule of law. This power to self-generate norms that escape public and private oversight underlines how code does not only play the role of the law, but it is also a source of law.

The expansion of the rule of tech defines another example of how the rule of law is not only connected to a certain territorial dimension but also to multiple spaces that shape each other in a process of mutual influence. The rule of tech does not develop its norms in an isolated framework but in a system of competing normative powers that shape how artificial intelligence systems are designed and implemented over time. The main concerns come from the opacity of this normative layer and the limit of oversight. If these technologies will increasingly develop norms that affect daily lives, the question for constitutional democracies remains how to ensure that the rule of law has enough space to protect constitutional values. Therefore, the challenge for constitutional democracies is how to limit the marginalization of the rule of law as driven by the consolidation of the rule of tech in the digital age.

The position of the rule of law in the algorithmic society is critical to mitigate the challenges raised by the normative power of artificial intelligence. Regulation can play a role in expanding the spaces for the rule of law, thus providing safeguards and remedies, particularly when the rule of tech intimately touch upon constitutional rights and democratic values. Among different approaches, including the implementation of adversarial systems to inject more checks and balances within automated decision-making or the participation of more stakeholders in the phase of design and monitoring, the introduction of

105. Artificial Intelligence Act, supra note 17, at 46–47.
procedural safeguards and remedies can provide a first answer to limit
the challenges raised by the rule of tech.

The launch of the Artificial Intelligence Act is a way to deal with the
marginalization of the rule of law. By relying on a risk-based approach,
this legal instrument recognizes that algorithmic systems cannot be
implemented in any case, thus limiting technological delegation. Even if
the Artificial Intelligence Act fails to provide a flexible framework and
remedies against the impact of different applications, it defines a first
step to address the expansion of the rule of tech, not only from a
European internal dimension but also with regard to external
interferences and influences by other areas on a global scale that would
affect the protection of European values.

The normative power of artificial intelligence raises questions for
digital constitutionalism to address the exercise of unaccountable
powers. The questions raised by the rule of tech are intimately
constitutional and lead to consider how far the rule of law can tolerate
the expansion of unaccountable forms of power in the algorithmic
society.
Algorithmic Constitutionalism

OREN PEREZ & NURIT WIMER

ABSTRACT

The increasing encroachment of artificial intelligence (AI) on social life raises various risks to society, most prominently in the info-spheres created and controlled by Google, Facebook, Apple, and Amazon. We examine these risks through an in-depth discussion of the Facebook content moderation regime, which is already partially controlled by algorithms. We argue that the idea of ethical engineering, which was developed in the literature as a solution to the challenge of governing AI, is inadequate for various reasons. We develop a different approach to coping with the risks of AI governance, which we have termed “algorithmic constitutionalism.” Our approach rests on three pillars: (a) layered architecture that consists of two levels of code: (i) operative or object level and (ii) metalevel, whose purpose is to shield the core principles of the system from algorithmic-initiated changes; (b) algorithmic metareasoning, which allows the system to operate simultaneously at the two levels, so that it can self-monitor, verify, and potentially correct in real-time operations at the object level if they depart from the principles protected by the metacode level; and (c) correction by deliberation. We elaborate the idea of algorithmic constitutionalism and demonstrate how it can be applied to the Facebook content moderation regime. As part of this elaboration, we also consider the tension between societal and algorithmic constitutionalism. Paradoxically, the attempt to subject the AI algorithm to external deliberative control also opens the door for the AI agent to intervene in that process, potentially undermining its very purpose. We conclude by exploring the implications of our argument for the new European Digital Services Act, which came into force in October 2022.
The increasing encroachment of artificial intelligence (AI) on social life raises various risks to human welfare and human rights. There is a growing literature arguing that a prepotent artificial intelligence (PAI) may present an existential risk to humanity. We argue that a focus on PAI shifts our attention away from the risks that existing hybrid human-digital systems present to human society today. These risks are most prominently found in the info-spheres created and controlled by Google, Facebook, Apple, and Amazon. Although these risks may not be existential, they are sufficiently serious to warrant close scrutiny. We examine these risks through an in-depth discussion of the Facebook governance system, which is already partially controlled by algorithms that perform various tasks, such as content moderation and classification, friend recommendation, ad delivery, and more. Second, we argue that ethical engineering, one of the primary approaches to the risk of PAI, does not offer a satisfactory solution to the challenge of governing AI. The first difficulty with ethical engineering concerns moral pluralism. Despite humanity’s long history of moral inquiry, we lack a firm understanding of what the “right” moral theory is. The problem of moral pluralism becomes more acute as the

1. We borrow the term from ANDREW CRITCH & DAVID KRUEGER, AI RESEARCH CONSIDERATIONS FOR HUMAN EXISTENTIAL SAFETY (ARCHES) 12–15 (2020). Other authors use the term “super-intelligent AI” or simply “super AI.” See MITJA KOVAČ, JUDGEMENT-PROOF ROBOTS AND ARTIFICIAL INTELLIGENCE: A COMPARATIVE LAW AND ECONOMICS APPROACH passim (2020).


range of problems AI is expected to resolve becomes broader. A second shortcoming of the ethical engineering approach is the fact that the interaction of an AI agent with the external environment and the knowledge it may acquire through that process, coupled with a capacity for self-renewal, may lead it to develop new objectives and new interpretations of existing ones, which may be inconsistent with the ethical principles that the human designer has intended to implant in it.

In this article, we develop a different approach to coping with the risks presented by algorithmic governance, which we have termed "algorithmic constitutionalism" (AC). Our approach rests on three pillars: (a) Layered architecture, which consists of two levels of code: (i) operative or object level and (ii) metalevel. The purpose of this layered architecture is to shield the core principles of the system (which are located at the metacode level) from algorithmically initiated changes that can radically transform the system. (b) Algorithmic metareasoning, which allows the system to operate simultaneously at the two levels, so it can self-monitor, verify, and potentially correct, in real time, operations at the object level if they depart from the principles protected by the metacode level. The metacode can be figuratively described, following Giacomo et al., as a "restraining bolt." Finally, (c) Correction by deliberation, which limits the ability of the metalevel code to initiate corrective actions by subjecting it to hardwired deliberation procedures (considering both the problem of moral pluralism and the risks that may emanate from an autonomous AI algorithm that has the power to rewrite the core principles of the system). As we elaborate below, this proposal implies a departure from the optimization principle that underlies the dominant approach in AI development. Our thesis builds on humanity’s long experience in taming power. Rather than relying on ethics and morality to guide it, human society has put its trust in politics, drawing on various decision-making procedures (from voting to polls), embedded in an intricate institutional structure of checks and

7. For example, the moral tradeoffs that underlie the moderation of speech on social media platforms are much more complex than those facing the control of autonomous vehicles. See Jean-François Bonnefon et al., The Trolley, the Bull Bar, and Why Engineers Should Care About the Ethics of Autonomous Cars, 107 PROCEEDINGS OF THE IEEE, 502, 504 (2019); Christoph Luetge, The German Ethics Code for Automated and Connected Driving, 30 PHIL. TECH. 547, 553 (2017).

8. Stefania Costantini, Ensuring Trustworthy and Ethical Behavior in Intelligent Logical Agents, 32 J. LOGIC & COMPUTATION 443, 444 (2022).

balances. Together, these three principles form the basis of what we call “algorithmic constitutional law.”

In the sections below we develop the idea of algorithmic constitutionalism and explore the legal, political, and technological challenges it faces while considering the connection between algorithmic and societal constitutionalism (SC). Our argument shares with SC the claim that state constitutions do not exhaust the universe of constitutionality, but it departs from SC by developing an algorithmic interpretation of constitutionality, which challenges the societal aspect of SC. We believe that one of the greatest challenges facing modern constitutional theory concerns the relation between algorithmic constitutionalism and SC. In the last two sections of the article, we further explore the tension between these two ideas.

To make the discussion more concrete, we examine a hypothetical scenario in which Facebook becomes fully controlled by a semi-PAI. The Facebook autonomous platform law, with its myriad of company statements, policies, internal standards, and algorithms, constitutes a perfect arena for analyzing the tension between algorithmic constitutionalism and SC. We consider in this context the recently approved EU Digital Services Act (DSA), which establishes a new regulatory regime for social media platforms, significantly reducing their capacity for self-regulation.

The paper proceeds as follows. Section A discusses the risks associated with PAI. Section B examines the intersection between human agency and AI in the governance of Facebook (focusing on the domain of content moderation). Section C discusses the risks of removing humans from the Facebook compliance system and making it entirely AI-controlled. Section D describes our concept of algorithmic constitutionalism and applies it to the realm of Facebook. Section E concludes with a discussion of algorithmic constitutionalism and SC, drawing also on the new regulatory regime of the DSA.

I. THE RISK OF PAI AND THE LIMITS OF ETHICAL ENGINEERING

The idea that a PAI system may present an existential risk to humanity has received significant attention in academic literature in the past few years. Nick Bostrom has defined the idea of PAI as “an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom and social skills.” He argued that if PAI evolved, humanity would most likely not be able to stop it.

Bostrom emphasized that PAI has several unique properties that make it distinctive compared to other novel technologies, including (a) autonomy manifest in the ability to make independent decisions and take initiatives; (b) the development of distinctive motivational tendencies (not shared by humans); (c) capacity to self-replicate and self-improve; and (d) unique cognitive architecture which may guard an AI agent against some kinds of human error and bias, while at the same time making it more susceptible to other types of mistakes. The primary source of risk underlying the emergence of PAI is that it may become misaligned with human interests and goals and, given the powers and capabilities of PAI, such misalignment can produce catastrophic consequences for humanity. Bostrom noted that we may not be able to predict with precision the moment when PAI will appear.

Critch and Krueger compared the potential effect of PAI to those of the agricultural and industrial revolutions, which have changed the course of humanity. They argued that the emergence of PAI is likely to follow a nonlinear path. AI power will grow steadily until one AI system reaches a threshold of self-improvement, at which point it will quickly outperform others by many orders of magnitude, leading to a radical transformation of the current world system. They compared the development of PAI to the spread of humanity on Earth. Just as the nonhuman world was unable to contend with the expansion of humans, humans will not be able to cope with the powers of PAI. The main risk, Critch and Krueger argued, is that once PAI technology is deployed it will not be possible to reverse or stop its transformative effect. They

16. *Id.* at 18.
17. *Id.* at 13.
noted that whereas machines can operate in a wide spectrum of physical environments, humans can survive only in a relatively narrow set of physical conditions and are therefore vulnerable to potential changes in the environment that may be beneficial to machines but not to humans.

The primary approach of the literature to this risk, which draws on the early writings of Isaac Asimov, is based on the idea of ethical engineering. According to this approach, as AI systems and robots become more advanced and intelligent, they must be subjected to ethical norms. It is useful in this context to distinguish between the ethics of AI and ethical AI. The ethics of AI deals with the principles that govern the development of AI, ensuring that ethical dilemmas about the interaction of AI with humans and nature are taken into account in the design phase. Ethical AI focuses on the AI system itself: how ethical principles can be incorporated into the AI system, ensuring that it operates ethically in its interaction with other agents (humans, animals, and other AI systems).

One of the best-known approaches for ethical AI is Isaac Asimov’s three laws of robotics, introduced in the 1950s:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.

2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.

3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

We argue that the approach of ethical engineering is problematic in two respects. First, given the condition of moral pluralism, it is difficult

18. Id. at 17.
19. Id.
to see how one can develop an algorithm that can perfectly reconcile between competing moral answers and practical dilemmas. This problem reflects the incommensurability of moral values, where values are considered incommensurate “if it is neither true that one is better than the other nor true that they are of equal value.”

Advocates of ethical engineering have not given sufficient attention to this problem. Thus, for example, the recent IEEE Standard Model Process for Addressing Ethical Concerns during System Design takes it for granted that values can be successfully prioritized through a process of elicitation and consultation with internal and external stakeholders and a comparison with external authoritative sources. Human societies have not resolved the problem of incommensurability by developing a universal ranking of values (metaethics) but by establishing an intricate network of political, executive, and judicial institutions, which, through a system of checks and balances, can resolve conflicts between incommensurables.

The second problem with the idea of ethical engineering concerns the risk that a PAI agent may modify its goals while developing its understanding of the world, making these goals inconsistent with those of its creators. Instilling the AI agent with a system of prioritized values does not provide a robust defense against the risk of misaligned AI creativity.

II. THE INTERSECTION BETWEEN HUMAN AGENCY AND AI IN FACEBOOK INTERNAL GOVERNANCE

The activity of Facebook users is subject to a variety of rules set and enforced by Meta. The main set of rules that governs Facebook users’ behavior is the “Community Standards.” That system is almost fully

25. Id. at 39; see also Sarah Spiekermann & Till Winkler, Value-Based Engineering with IEEE 7000, 41(3) IEEE TECH. AND SOCY MAGAZINE 71, 75 (2022).
26. See generally Timothy Endicott, Proportionality and Incommensurability, 40/2012 OXFORD LEGAL STUDIES RESEARCH PAPER SERIES (2013) (exploring the frameworks through which such incommensurables are resolved).
27. Wolfhart Totschnig, Fully Autonomous AI, 26 SCI. & ENG’G ETHICS 2473, 2473 (2020); see also Mitja Kovac, Autonomous Artificial Intelligence and Uncontemplated Hazards: Towards the Optimal Regulatory Framework, 13 EUR. J. RISK REGUL. 94, 95–96 (2022).
28. Alongside the Community Standards, there are other policies this article will not cover, such as the “Page, Groups and Events” policy “that app[lies] to all creators and [managers] of [p]ages, [g]roups, and [e]vents on Facebook, the “Branded Content Polic[y]”
controlled by Meta, with limited public involvement. Our goal in this section is to expose the intersection between AI tools and human agency in the governance of Facebook’s internal platform law. The involvement of algorithms on the Facebook platform extends beyond the enforcement of the community standards (e.g., ad delivery). However, we have limited our analysis to the role of AI tools in the implementation of the community standards because we are using it primarily to illustrate our broader argument about algorithmic constitutionalism. Our analysis relies on publicly available data that probably do not provide a complete picture of the involvement of AI in Facebook compliance processes, but this informational gap is not critical for our general argument.

A. Facebook Community Standards

Facebook Community Standards govern the behavior of users with respect to content. The standards have been created by Meta’s human team, with input from the public and experts in diverse fields, such as technology, public safety, and human rights. The purpose of the standards is to, on one hand, create a place for expression and give people a voice, and, on the other, to prevent the publication of content that may undermine one of the following objectives: authenticity, safety, privacy, or dignity. Content that violates one of these objectives can still get published if it is newsworthy and in the public interest. Such a determination is made through a formula that balances the damage that may be caused by the publication with public interest in the

---

29. See, e.g., Ali et al., supra note 4; J. Bobadilla et al., Recommender Systems Survey, 46 KNOWLEDGE-BASED SYSTEMS 109, 109 (2013); Ash et al., supra note 4, at 14–16; DEL VIGNA ET AL., Hate Me, Hate Me Not: Hate Speech Detection on Facebook (2017); Shangrong Huang et al., Social Friend Recommendation Based on Multiple Network Correlation, 18 IEEE TRANSACTIONS MULTIMEDIA 287, 287 (2016); MacAvaney et al., supra note 4, at 6.

30. For example, in a study conducted by Kate Klonick, significant gaps were discovered between the declared community standards and the rules according to which the Facebook review teams intervened in content. See Kate Klonick, The New Governors: The People, Rules, and Processes Governing Online Speech, 131 HARV. L. REV. 1598, 1648, 1669–70 (2018).


32. Id.
content.\textsuperscript{33}

The standards are divided into six main sections dealing with: violence and criminal behavior, safety, objectionable content, integrity and authenticity, respect for intellectual property, and content-related requests and decisions (see Table 1 below for details). For each section, the standards elaborate the differences between various categories of content according to the following schema\textsuperscript{34}:

- Content that is not allowed, and
- Content that requires additional information or context, content that is allowed with a warning screen, or content that is allowed but can be viewed only by adults aged eighteen and older.

For example, in the section on violence and incitement, content that users cannot post and that will be removed includes: “[s]tatements of intent to commit high-severity violence,” “[c]alls for high-severity violence;” and “[a]spirational or conditional statements to commit high-severity violence.” Some threatening language, however, requires additional information or context for Meta to act upon, such as “[v]iolent threats against law enforcement officers” or “[v]iolent threats against people accused of a crime.”\textsuperscript{35} Table 1 summarizes the content categories governed by the community standards.

---


\textsuperscript{34} See Facebook Community Standards, supra note 28.

Table 1: Content categories covered by the community standards

B. The Hybrid (AI/Human) Facebook Internal Compliance System

The community standards are enforced by a hybrid compliance system that combines AI algorithms and human agents (users’ input, review teams, and external experts). The enforcement process consists of four stages (Figure 1):

1. Preliminary *(prima facie)* flagging of content as breaching the community standards.
2. Investigation of the preliminary flagging and a determination whether a violation has in fact occurred.

3. Imposition of sanction or other administrative response.

4. Reexamination and appeal concerning the punitive action.

In addition to these four steps, Meta has employed a system called cross-check in the past few years, which provides some categories of entities and content, special consideration, and is governed by a distinct set of rules. Below we briefly describe how AI and human agents intersect in the enforcement process.

1. **Preliminary (Prima Facie) Finding of Violating Content**

   The enforcement process is initiated by flagging an item of content as potentially violating. Such flagging can occur either through user reports or through AI detection, based on Natural Language Processing technology. To enhance its capacity to detect violating content, Meta developed an AI model called “Few-Shot Learner” (FSL). Traditional AI models require a massive amount of labeled data points to train the model. Collection and labeling of these data points can take several months. Under FSL, violating content can be detected even without the availability of a large body of labeled data. FSL can respond to new types of harmful content within weeks instead of months. It can work in more than one hundred languages and is able to learn from various types of data, including images and text. Meta describes this process as follows:

   “It’s first trained on billions of generic and open-source language examples. Then, the AI system is trained with policy-violating content and borderline content [that were] labeled over the years. Finally, it’s trained on

---

39. Id.
condensed text explaining a new policy . . . the key idea is to convert the class label into a natural language sentence which can be used to describe the label, and determine if the example entails the label description.”  

FSL can also anticipate new types of policy violations. This information can allow Meta to

“understand the sensibility and validity of the proposed definitions. It casts a wider net, surfacing more types of ‘almost’ content violations that policy teams should know about when deciding or formulating at-scale guidance for annotators who train new classifiers and the human reviewers helping to keep our platform safe . . . . Our long-term vision is to achieve human-like learning flexibility and efficiency, making [its] integrity systems even faster and easier to train and better able to work with new information.”

2. Investigation of a Prima Facie Flagging and Determination of Violation

After content has been flagged as potentially in violation, a determination must be made whether it violates the community standards. This determination can be made either by AI algorithm or a review team. Usually, content flagged (by AI or by user reports) as potentially violating is submitted for review by the AI tool to determine whether it is indeed violating. There are five potential responses:

1. There is a violation.
2. There is no violation.
3. There is no violation, but the content is sensitive or misleading.
4. There is no violation, but the content is problematic.
5. Unable to decide if there is a violation.

40. Id.
41. Id.
If it is determined that there was no violation, the suspicious content is “acquitted,” and the inquiry is concluded. If the content is found to be violating, a decision is required regarding the appropriate sanction. If the content is determined to be sensitive or misleading (but not violating), it will usually not be removed, and less severe actions aimed at protecting other users will be taken. Content is classified as problematic when it does not violate community standards but may nevertheless harm users and impair their experience of using Facebook,\textsuperscript{42} for example, when the content contains links to sites that are packed with advertisements, or when it leads to shocking, disturbing, or malicious advertisements.\textsuperscript{43} The same is true of content that constitutes “clickbait,”\textsuperscript{44} that is, content that includes misleading, sensational, or spammy stories whose headlines are intended to grab users’ attention in order to click on a link.

When the AI algorithm fails to decide whether there has been a violation, the content is passed on to human review teams for determination. This happens in borderline cases, where the wording is vague or the context is important for reaching a decision.\textsuperscript{45} The review teams consist of trained human examiners who have expertise in the domain of the suspected violation and in the language in which the content was published. The review teams intervene in the decision-making process in the following cases: (a) when the AI algorithm has been unable to determine whether there is a violation; (b) when the suspicious content has been flagged by users; (c) in content categories that were predefined by Meta as necessitating human review, for example, content that requires attention to subtle nuances (such as bullying);\textsuperscript{46} or (d) content that falls under Meta’s cross-check system (see below for extended elaboration). The order through which the review teams consider the flagged posts is determined by AI. The AI algorithm considers three factors in deciding how to prioritize the flagged content: (a) the likelihood that it violates community standards, (b) the


\textsuperscript{45} See How Meta Prioritizes Content for Review, supra note 37.

probability that its publication will cause harm, and (c) its virality.47

The review teams can choose between the five responses described above. If the review team is unable to decide whether a violation has occurred, it can consult with specialized experts from the Global Operations or Content Policy Teams.48 Meta notes that it uses the information obtained from the decisions of the human review teams to train and improve the accuracy of its AI algorithm.49

a. The Cross-Check System

The existence of the cross-check system was exposed by the Wall Street Journal in September 2021. The report, which was based on documentation produced by former employee and company critic Frances Haugen, described the system as exempting Meta’s most influential users from normal content moderation processes.50 The Oversight Board received a request from Meta to review the system and consider how it could be improved.51 The following analysis is based on the board’s advisory opinion, published on December 6, 2022.52 Meta has not been transparent about the system, and the system’s details have become public only after the publication of the board’s recommendations.53 The cross-check system provides additional layers of review for two types of predefined content in an effort prevent overenforcement or false positives. First, cross-check provides guaranteed additional human review of content created by specific entities.54 Under this scheme, which is called Early Response Secondary Review (ERSR), if a protected entity posts content that is identified as violating, it will not be removed according to the procedures that apply

50. OVERSIGHT BOARD, POLICY ADVISORY OPINION ON META’S CROSS-CHECK PROGRAM 6 (December 2022).
51. Id. at 3.
53. See OVERSIGHT BOARD, supra note 50, at 4.
54. See OVERSIGHT BOARD, supra note 50, at 9–15. Meta reported to the Oversight Board that it includes entities on ERSR lists according to the following categories: “(1) civic and government; (2) significant world events; (3) media organizations, businesses, communities and creators, including advertisers; (4) historically over-enforced; (5) legal and regulatory or entities for which erroneous action may present legal risk to Meta, for example in the context of ongoing litigation; (6) entities whose content is under review, meaning cases where action by any reviewer would undermine ongoing deliberation or would present risk to Meta"
to regular users but instead will be sent for extra levels of review. The second scheme that forms part of the cross-check system is General Secondary Review (GSR). Whereas ERSR applies to all content posted by specific entitled entities, GSR may apply to any content posted on the platform, based on a determination by the cross-check ranker algorithm, that it merits additional review because of a high likelihood of being a false positive. The cross-check ranker algorithm relies on the following features: ‘topic sensitivity (how trending/sensitive the topic is), enforcement severity (the severity of the potential enforcement action), false positive probability, predicted reach, and entity sensitivity.’

In both categories, the cross-check process kicks in after a piece of content is flagged for review by AI or is reported by a human user and before the imposition of any sanction. The additional review is performed by human moderators. One of the issues that was criticized by the board was the unequal treatment Meta gives content in the ERSR and GSR streams: whereas content that is reviewed under the ERSR benefits from up to four review cycles, content that is reviewed under the GSR benefits from a maximum of two review cycles. The enforcement actions against content recognized as eligible for cross-check is suspended until the end of the review process. Under the current cross-check system, content reviewed under ERSR receives priority over GSR content. Due to the limited capacity of the Early Response teams, not all GSR content is reviewed.

The Oversight Board exposed another system that blocks some enforcement actions outside of the cross-check system, which Meta referred to as “technical corrections.” “Technical corrections” are automatic exceptions to content policy enforcement. As explained by Meta, a “technical correction” applies only to a specific entity for a specific policy violation and does not serve to bar enforcement for other policy violations. At the time of the board’s briefings with Meta, it stated that it applied about a thousand technical corrections per day. Meta did not disclose how many and what type of entities have benefited from a

55. Id. at 11.
56. Id. at 18.
57. Id.
58. Id. at 19.
59. Id. at 14, 21.
60. Id.
61. Id.
62. Id. at 19–20. Unlike ERSR, content in the GSR is put in a queue pending review. Content that is not reviewed within two to four days is popped out of the GSR queue. When this happens, Meta reverts to its initial enforcement decision (e.g., removal or a warning screen) without further review. Id. at 20.
The Oversight Board has strongly criticized the cross-check program, highlighting its unequal nature (which was manifested in the privileged treatment of business partners and the lack of proper protection to entities that are likely to produce expressions with significant human rights value) and its nontransparent functioning (e.g., Meta’s refusal to share data regarding the identity of the business partners enjoying cross-check privileges).64

3. Compliance Action

Once a decision has been made whether an item of content violates community standards, compliance action is taken against the user. A decision on that action can be made by either a review team or by an AI algorithm.65 The nature of the sanction or other action will be determined based on the circumstances and by the nature of the “offense.” There is a broad spectrum of potential compliance responses with varying levels of severity:

- Removing the infringing content.
- Limiting the exposure of the content by reducing its News Feed ranking (e.g., in the context of false information).
- Restricting the visibility of content for some users, for example, those under eighteen (e.g., in the case of content related to alcohol, tobacco, sex toys, or weapons).
- Adding a warning screen/tag (e.g., in relation to content

63. Id. at 22.
64. Id. The board provided Meta with a series of detailed recommendations. Id. at 35–48.
65. Meta states that it uses a dedicated enforcement technology system that determines whether to take action, but does not specify in which cases decisions are made by algorithms and in which they are made by human teams. How Enforcement Technology Works, META: TRANSPARENCY CENTER (Jan. 19, 2022), https://transparency.fb.com/enforcement/detecting-violations/how-enforcement-technology-works/; see also ÁNGEL DÍAZ & LAURA HECHT-FELELLA, DOUBLE STANDARDS IN SOCIAL MEDIA CONTENT MODERATION 10 (2021).
that refers to nonconsensual sex).\textsuperscript{66}

- Disabling the account in cases of serious violations, such as the publication of content that contains child pornography or physical abuse.\textsuperscript{67}

- Reporting to relevant law enforcement authorities.\textsuperscript{68}

When users are involved in repeated violations, they may be subject to additional sanctions, such as restricting accounts for a limited period or, in severe cases, disabling the account.\textsuperscript{69} Meta uses the method of counting strikes in determining the severity of the measures it takes against users. The system keeps a record of the violations (strikes) of each user. If the number of strikes exceeds a certain threshold, it may lead to the restriction or disabling of the account.\textsuperscript{70} Users cannot post content or create pages using a restricted account.\textsuperscript{71}

The first strike usually results in a warning with no further action. Two strikes may lead to restrictions limited in time (e.g., one day), three strikes to three days, and four strikes to seven days. From the fifth strike onward, the account may be limited for thirty days for each further strike\textsuperscript{72} and may be entirely disabled, depending on the severity of the violation.\textsuperscript{73} There may be cases of serious violations in which an account is disabled after one strike, as in the case of posting child pornography. Meta disables certain types of accounts as soon as it notices their existence, such as accounts of convicted sex offenders or accounts that display false identities.\textsuperscript{74}

The rules applicable to groups and to pages that consistently violate

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{68} Id.
\item \textsuperscript{69} \textit{Restricting Accounts}, META: TRANSPARENCY CENTER (Feb. 23, 2023), https://transparency.fb.com/he-il/enforcement/taking-action/restricting-accounts/.
\item \textsuperscript{71} \textit{Restricting Accounts, supra} note 67.
\item \textsuperscript{72} Id.
\item \textsuperscript{73} \textit{Disabling Accounts}, META: TRANSPARENCY CENTER (Jan. 19, 2022), https://transparency.fb.com/he-il/enforcement/taking-action/disabling-accounts/.
\item \textsuperscript{74} Id.
\end{itemize}
\end{footnotesize}
community standards differ from those that apply to private users. For example, Meta may require page and group managers to approve each post before it is published. During civil unrest, Meta may restrict accounts of public figures for longer periods if it finds that they incite violence.

Meta notifies users when they post content that violates community standards. The notice of infringement usually includes a mention of the standard that was violated and a brief description of the reason why the content was removed. When an account is disabled, Meta notifies the user when trying to log into the account. Meta also notifies the user when it issues a strike and when the account has been restricted.

4. Appeal and Reexamination Procedures

If users believe that they have been subject to a compliance action for no justified reason and that the content that they shared does not violate Facebook’s community standards, they can request that the post be reexamined. The reexamination process usually takes up to twenty-four hours. If the reexamination reveals that there was a mistake and the content was removed or concealed without a valid reason, the content is restored and the user is notified. The review is conducted by human audit teams.

Users who do not agree with the outcome of the reexamination process can appeal to the Oversight Board within fifteen days from the day on which they received the final decision. The board reviews decisions by Facebook and Instagram as well as cases referred to it by

---

76. Restricting Accounts, supra note 67.
79. Disabling Accounts, supra note 71.
81. I Don’t Think Facebook Should Have Taken Down My Post, FACEBOOK: HELP CENTER, https://www.facebook.com/help/2090856331203011 (last visited Mar. 6, 2023) (demonstrating that Facebook does not specify in which cases the right to have the action reexamined is not granted).
82. OVERSIGHT BOARD, OVERSIGHT BOARD BYLAWS 15, 22 (2022).
83. I Don’t Think Facebook Should Have Taken Down My Post, supra note 64.
Meta.\footnote{Oversight Board, Oversight Board Bylaws 15, 22, 27 (2022).} The Oversight Board has the right to choose which cases it reviews.\footnote{Id. at 15.} Its decisions are binding, and Meta is committed to implementing them.\footnote{Id. at 15.} The board is overseen by a group of trustees.

According to its bylaws, the Oversight Board must set up two subcommittees, the Membership Committee and the Case Selection Committee, and may establish additional ones. Members of the Case Selection Committee serve for three months and set the criteria for prioritizing cases. The committee reaches its decisions by majority rule but can be overruled by a majority vote of the full board.\footnote{Id. at 8–9.} The Oversight Board can review only a very small portion of the total number of Facebook content moderation decisions. According to Facebook bylaws, the review process should last ninety days.\footnote{See id. at 15.} Each case is preliminarily reviewed by a panel of five members that can request more information from Facebook and other sources, such as external experts or advocacy groups, that reflect a range of perspectives. The panel receives a statement from the user who submitted the request, the case history (delivered by Meta), relevant policy (also from Meta), and any additional external information that any member of the panel may ask for.\footnote{Id. at 16.}

Meta is committed to applying board decisions beyond the case being reexamined. For example, if the Board decides that certain content violates community standards,\footnote{See generally id. at 26 (including a violation of “Community Standards” as a requirement for board review).} Meta will conduct a search aimed at identifying similar content and, if it is technically feasible, will take measures against it.\footnote{Id. at 25.}
Figure 1: Description of Facebook’s content moderation framework
III. THE RISKS OF A SHIFT TO FULL AI GOVERNANCE: AN IMAGINARY FACEBOOK WITH NO HUMANS IN THE LOOP

Facebook’s compliance framework is a hybrid system in which human agents interact with AI algorithms. In this section, we explore the potential risks of an imaginary scenario in which Facebook becomes governed entirely by AI and humans cease to be part of the compliance loop. We assume that this AI system possesses part of the qualities of a PAI, in particular, a capacity to make autonomous decisions that may be misaligned with the objectives of its controllers. A shift to full AI governance raises two types of risks: potential increase in cases of over and under enforcement (false positives and false negatives) due to loss of human oversight and algorithmic bias; and potential misaligned, AI-driven revision of the system’s fundamental normative structure. This hypothetical scenario seems like a natural extension of existing practices on social media platforms, which already use algorithms extensively to manage their operations. This reliance allows them to scale up their operations and cut costs.

One of the key implications of removing humans from the compliance loop is the loss of the distinction between preliminary, first-order, compliance actions and second-order, review/appeal decisions. Most of the preliminary compliance decisions at Facebook are made by automated means. Human involvement in the preliminary stage is limited to (a) difficult or borderline cases and (b) failure of the AI algorithm to decide whether certain content violates the standards.93 The second-order review stage is conducted entirely by humans. If human agents are no longer involved in the compliance process, the distinction between the two stages of decision will disappear, resulting in an integrated, automated compliance process.

The question is whether a shift to a fully automated compliance process would lead to an increase in the number of false positives and false negatives, in a way which adversely affects users’ welfare and rights. In their recent book, Noise: A Flaw in Human Judgment, Kahneman, Sibony, and Sunstein argue that algorithmic prediction techniques represent “significant improvements on human judgment. The combination of personal patterns and occasion noise weighs so heavily on the quality of human judgment that simplicity and noiselessness are sizable advantages.”94 As Jhaver and colleagues noted

93. See supra Section II (ii).
94. DANIEL KAHNEMAN ET AL., NOISE: A FLAW IN HUMAN JUDGMENT 133 (2021). We must distinguish between noise, which “consists of unwanted variability in judgments,” and bias, which “consists of any systematic error that inclines people’s judgments in a
in a study about content moderation on Reddit, “[a]utomating moderation not only facilitates scalability, it also enables consistency in moderation decisions.”

Other works, however, have offered an opposing perspective, which emphasizes the risk of algorithmic bias, which occurs when “the outputs of an algorithm benefit or disadvantage certain individuals or groups more than others without a justified reason for such unequal impacts.” Such bias can have several causes: faulty design, a training dataset that is contaminated with social biases (either owing to faulty sampling or because of biases entrenched in society), or bias-reinforcing loops. Researchers and activists have pointed out algorithmic bias in a variety of environments, including criminal justice, healthcare, hiring, and credit scores.

Meta has been involved in several cases of algorithmic bias, which demonstrates that the foregoing concerns are not unfounded. Multiple studies have shown that the Facebook ad delivery system can be skewed by gender or race because of hidden algorithmic optimization, even when not requested by the advertisers. A recent lawsuit filed against Meta by the US Department of Justice alleged that housing advertisements on Meta platforms have discriminated against users based on their race, color, religion, sex, disability, familial status, and national origin, in violation of the Fair Housing Act. As part of a settlement agreement, which was signed in June 2022, Meta agreed to stop using the advertising tool known as the “Special Ad Audience” that had a discriminatory effect. Meta has also agreed to develop a new particular direction.” Cass R. Sunstein, Governing by Algorithm? No Noise and (Potentially) Less Bias, 71 DUKE L.J. 1175, 1178 (2022).


97. See id. at 390; see also Ryan S. Baker & Aaron Hawn, Algorithmic Bias in Education, 32 INT’L J.A.I. EDUC. 1052, 1057–61 (2022) (discussing forms and sources of bias in educational technologies and algorithms); BRIAN CHRISTIAN, THE ALIGNMENT PROBLEM: MACHINE LEARNING AND HUMAN VALUES 33–34 (2020) (“Bias in machine-learning systems is often a direct result of the data on which the systems are trained—making it incredibly important to understand who is represented in those datasets, and to what degree, before using them to train systems that will affect real people.”).

98. See Kordzadeh & Ghasemaghaei, supra note 78, at 390, 392.

system to address racial and other disparities caused by its use of 
personalization algorithms in its ad delivery system for housing ads.\textsuperscript{100} 
A recent study published by the NYU Brennan Center for Justice has 
shown that the Facebook content moderation approach to hate speech 
and violent extremism content was biased against minority groups.\textsuperscript{101}

The problem with removing human agency from the Facebook 
compliance system is that there is no general algorithmic debiasing tool 
that can reliably replace human oversight.\textsuperscript{102} This reflects the 
complexity of algorithmic bias, its diverse manifestations, and multiple 
potential causes. Making the Facebook compliance system more 
consistent and noiseless runs the risk of uncontrolled algorithmic bias.

The second type of risk that may result from entrusting the 
governance of Facebook to an AI system emanates from its semi-
prepotent capacities: its ability to initiate, on its own “volition,” a 
process leading to the revision of Facebook’s fundamental norms in a 
direction that may be incompatible with fundamental human rights. 
This risk has several aspects that should be distinguished. The first 
aspect focuses on the potential adverse effects of an overzealous, 
unreflective optimizing AI agent. Such an agent could pursue its own 
interpretation of an underspecified objective or rule and the means for 
realizing it, in ways and scales that were not foreseen by the designer. 
The second aspect concerns the risk that an autonomous AI agent may 
develop novel objectives that are distinct (and potentially misaligned) 
from those that were given to it by its designer.\textsuperscript{103}

\textsuperscript{100} See Settlement Agreement & Final Judgment at 2, 6–7, United States v. Meta 
\textsuperscript{101} See Díaz & Hecht-Felella, supra note 48, at 5–6, 10–11. 
\textsuperscript{102} See Jhaver et al., supra note 77, at 31:26–31:28. See generally Andreas Tsamados 
et al., The Ethics of Algorithms: Key Problems and Solutions, 37 A.I. & Soc'y 215 (2022) 
(reviewing relevant literature and discussing six ethical issues derived from algorithm 
use); Manish Raghavan et al., Mitigating Bias in Algorithmic Hiring: Evaluating Claims 
and Practices, 2020 PROC. OF 2020 CONF. ON FAIRNESS, ACCOUNTABILITY & TRANSPARENCY 
469, 475–78 (analyzing the use and challenges of algorithmic de-biasing in the context of 
disparate impact hiring practices); Jason R. Bent, Is Algorithmic Affirmative Action 
Legal?, 108 GEO. L.J. 803, 809–24 (2019) (discussing causes of algorithmic bias, 
ineffectiveness of algorithm blindness, and tradeoffs between fairness and accuracy). 
\textsuperscript{103} See Adrien Ecoffet et al., Open Questions in Creating Safe Open-Ended AI: Tensions 
Between Control and Creativity, 2020 CONF. ON ARTIFICIAL LIFE 27, 28–30; see also 
Bostrom, Ethical Issues in Advanced Artificial Intelligence, supra note 14, at 279 
(discussing the possibility for superintelligence, which could have different motives than 
humans, to develop its own plans and goals). In a recent and interesting paper, Müller 
and Cannon argue that it is unplausible to attribute an unreflective-optimization risk to 
general intelligence, since a self-reflection capacity is expected in such intelligence. See 
Vincent C. Müller & Michael Cannon, Existential Risk from AI and Orthogonality: Can We 
Have It Both Ways?, 35 RATIO 30 (2022). What we should worry about, they argue, is the
We can illustrate these risks by considering the Facebook mission statement: “to give people the power to build community and bring the world closer together.” Facebook has changed its mission statement several times since it was established. The last change was made in 2017. The Facebook mission statement uses broad and vague language, and it is easy to imagine how an attempt to optimize the mission can generate results that would be considered morally problematic. The mission statement does not specify the types of communities it seeks to support or the means that can be used to bring people together. How will our imagined AI controller respond to an emerging community of neo-Nazis or ISIS supporters? What strategy will it adopt if it finds that fake news regarding climate change or vaccines can be effective in bringing people together? These are not hypothetical scenarios, as Facebook has already been implicated in facilitating the dissemination of fake news regarding the COVID-19 epidemic and climate change. A recent study by the Institute for Strategic Dialogue has found that Holocaust denial content has been actively recommended by the Facebook algorithm and that when a user follows public pages containing Holocaust denial content Facebook actively promotes further Holocaust denial content to that user.

Our optimizing AI controller may even go further in seeking to expand Facebook communities by abolishing rules that in its view constitute an impediment to such expansion, such as the rule that prohibits hate

(significant) damage that a “highly capable instrumental AI [not a PAI] can cause[] if designed or used badly.” Id. at 34.


105. See Nicholas A. John, Social Media Bullshit: What We Don’t Know About Facebook.com/Peace and Why We Should Care, 5 SOC. MEDIA + SOC’Y 1, 13 (2019). The previous version of Facebook’s mission statement was “Facebook gives people the power to share and make the world more open and connected.” Id. (emphasis added).

106. See id. Mark Zuckerberg published on his personal Facebook page a post in which he announced the decision to change Facebook’s mission statement and the reasons that led to this change. See Bringing the World Closer Together, FACEBOOK, https://www.facebook.com/notes/393134628500376/ (Mar. 15, 2021).

107. See generally Stephanie Preston et al., Detecting Fake News on Facebook: The Role of Emotional Intelligence, P.L.O.S. ONE, Mar. 2021, at 1, 4, 6–10, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0246757 (measuring the ability to detect fake news in Facebook posts, one of which contained fake news regarding climate change, among participants with varying levels of emotional intelligence and educational attainment).


108.

109. Id. at 1.
speech or the policy that reduces the visibility of false news. The problem may become even more acute if the AI agent revises Facebook’s overarching mission in a way that further supports such perverse expansion of communitarian communication.

IV. ALGORITHMIC CONSTITUTIONALISM

In this section we introduce and develop the concept of algorithmic constitutionalism and elaborate on the ways in which it can be implemented on a social media platform such as Facebook. The argument that follows is highly speculative, leaving many technical questions open. Many of the technological details underlying our argument are still at an early stage of development. The concept of algorithmic constitutionalism is based on three pillars: (a) Layered architecture, which aims to insulate the core principles of the system, located at the metacode level, from self-initiated changes that can modify the fundamental features of the system. We conceive the system as being comprised of two levels of code: (i) operative or object level and (ii) metalevel. The second pillar, (b) algorithmic metareasoning, which conceptualizes the system as operating simultaneously at two levels, allowing it to self-monitor, verify, and potentially correct operations at the object level in real-time if they depart from the principles protected by the metalevel code. We figuratively described the metalevel code as a restraining bolt. The final pillar, (c) correction by deliberation, which would limit the ability of the metalevel code to initiate corrective decisions that affect the fundamental features of the system and resort instead to deliberation as a mechanism for making such decisions. It is at this junction that societal and algorithmic constitutionalism can collide. Algorithmic constitutionalism thus establishes a hybrid framework in which human agents and algorithms collaborate in a way which is critical for the platform’s safety, legitimacy, and robustness.

The concept of algorithmic constitutionalism has some affinities with Stuart Russell’s principles of beneficial AI, which are:

1. The only objective of the machine is to maximize the realization of human preferences.

---


110. See generally False News, supra note 5, and False News Transparency, supra note 5 (describing Meta’s policies to reduce the distribution of false news and misinformation).

111. Costantini, supra note 8, at 445. See generally De Giacomo et al., supra note 9.

112. See Jhaver et al., supra note 77, at 27–28 (emphasizing the advantages of mixed-initiative regulation systems where humans work alongside automated systems).

113. STUART, supra note 23, at 173.
2. The machine is initially uncertain about what those preferences are.

3. The ultimate source of information about human preferences is human behavior.

Consistent with Russell’s argument, we want to instill in the AI governance system uncertainty about human preferences, which consequentially would be reflected in uncertainty about the correct or optimal interpretation of the rules (in our case, Facebook community standards) and of the overarching mission of the system. As Russell has argued, we want to design a humble AI that is willing to defer to humans:

A machine that assumes it knows the true objective perfectly will pursue it single-mindedly. It will never ask whether some course of action is OK, because it already knows it’s an optimal solution for the objective. It will ignore humans jumping up and down screaming, “Stop, you’re going to destroy the world!” because those are just words. Assuming perfect knowledge of the objective decouples the machine from the human: what the human does no longer matters, because the machine knows the goal and pursues it. On the other hand, a machine that is uncertain about the true objective will exhibit a kind of humility: it will, for example, defer to humans and allow itself to be switched off.114

In contrast to Russell, we reject the idea that the ultimate source of information about human preferences should be human behavior. Rather, we argue that the system should have a variety of procedures embedded in its algorithmic structure by which it can receive input from human agents (users) about the application of the standards, the system’s fundamental characteristics, and the potential revisions of rules.

Our approach also differs from Russell’s approach in that we reject the idea that the AI system should seek to optimize some measure of human preferences. Optimization coupled with superintelligence is a recipe for trouble, as Russell, Bostrom, and others have

114. See id. at 175.
demonstrated.\textsuperscript{115} The risk is most severe when relentless optimization comes with a limited capacity for self-reflection and incomplete rules. Optimization was never adopted as an overarching principle for the governance of human societies. Democratic societies are governed by the somewhat arbitrary rule of politics, which ultimately settles disagreement by a binary decision. This inherent arbitrariness is countered by a system of checks and balances that guarantees the reflexivity of the system, prevents it from becoming captured by a single voice, and allows an independent judiciary to play a crucial role.\textsuperscript{116}

Instead of optimizing, we propose the idea of political satisficing. Herbert Simon introduced the idea of satisficing. He argued that individuals do not choose alternatives as if they were maximizing (e.g., a utility function) but rather act as satisficers, decision makers who consider a limited number of alternatives investing limited cognitive effort, until they find one that is “good enough,” at which point they stop searching. A politically satisficing AI agent is one whose objective is not to maximize human preferences but rather to satisfice them subject to deliberative constraints, which are used to identify human preferences and resolve ambiguities about them.

The challenge is to elaborate how this abstract framework can be implemented in practice. The first step is to distinguish between the operative (base) level, which covers the implementation of Facebook’s community standards (or those of a similar social media platform), and the metalevel. The object-level AI agent has broad leeway for managing and improving the application of the standards but is not permitted to explicitly change the standards’ wording or revise the procedural principles they establish (e.g., right to appeal). Thus, the object-level algorithm is free to try to improve its classification of hate speech and violent extremism content, for example, by reducing the percentage of false positives (i.e., cases wrongly classified as violating the standards) and by improving the sensitivity of the system to social contexts (e.g., to the unique discourse of minority communities).\textsuperscript{117} But the object-level AI cannot revise the rules themselves (e.g., abolish the policy dealing with hate speech).

The metalevel code acts as a guardian of the community standards, which encode the core values of the community. The metalevel code also


\textsuperscript{117} See Jhaver et al., \textit{supra} note 77, at 24; Inioluwa Deborah Raji et al., \textit{The Fallacy of AI Functionality}, 2022 ACM CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY 959, 959 (2022); Diaz et al., \textit{supra} note 48, at 20–23 (providing critique).
governs how the standards may be revised. At the constitutional level, we usually distinguish between secondary or revision rules and first-order rules. Revision rules determine how other rules (both first- and second-order) can be revised. Because the metalevel code serves as the guardian of the algorithmic constitution, its capacity to directly revise the rules should be constrained.

We propose that the meta-AI agent operates in two modes of reflexivity: self- and user-initiated. At the self-initiated level, the meta-agent monitors the behavior of the object-level agent, identifies violations of the rules, and intervenes with corrective action when needed. The meta-agent does not monitor the entire behavior-state of the compliance system. Instead, it focuses only on the aspects that are critical for keeping the constitutional properties of the system intact, (e.g., fundamental procedural rules) such as the right of users to be notified of the imposition of sanctions and to appeal a decision that harmed their interests. Although the object-level agent cannot change the community standards explicitly, it can do so implicitly, for example, by refraining from notifying users that they were sanctioned or by blocking their attempts to send review requests. The cross-check system provides another illustration of how the object-level agent could attempt to circumvent the system’s formal rules through the creation of a list of privileged entities. When a violation is detected, the meta-agent can provide an immediate remedy, whether through a particularistic intervention (e.g., allowing a user to submit a review request) or by revising any changes to the object-level code that led to the problem (e.g., abolishing the cross-check bypass). This constitutional metamonitoring requires the system designers to assemble a limited set of “fundamental violating trigger events” and to incorporate them into the metalevel code. We also assume that the meta-AI agent is not limited to the initial list and can revise it through a recursive learning process in a way that should allow it to cope with unforeseen loops that may undermine users’ rights.

In articulating the idea of user reflexivity, we face the challenge of how to embed it into the metalevel algorithm. The first challenge concerns the design of the review process. As we elaborated above, in a fully AI-controlled environment, the preliminary enforcement phase and the secondary review process are conflated. We criticized this situation, noting the risks of algorithmic bias and unchecked transformative AI power. Responding to these problems requires developing a mechanism of review independent of the AI algorithm that controls the base level

---

118. See Costantini, supra note 8, at 445.
119. See Klonick, supra note 52.
compliance process. The second challenge concerns the need to create reflexive mechanisms that would enable a revision of the community standards in response to systemic problems that may emerge through the review and appeal process or external pressures.

The question is how to embed these reflexive capacities in the metalevel algorithm without endowing it with too much power. As we argued above, our approach is guided by two fundamental principles: (a) uncertainty about human preferences and, consequently, about the system’s overarching mission and the interpretation of the rules; and (b) the ultimate source of information about human preferences is not human behavior but human choice, as revealed through deliberation. These two principles should be algorithmically incorporated in a deliberative framework that enables the metalevel AI agent to realize the two modes of reflexivity.

There are two basic options for incorporating deliberation into the metalevel code. The first is to encode a dependence on the input of an external body, requiring the AI agent to turn to it in a predetermined list of cases. Such entities may include, for example, the Facebook Oversight Board, an out-of-court dispute settlement body certified under Article 21 of the DSA, or a multistakeholder entity, such as the Internet Governance Forum (IGF). Embedding in the code a linkage to such independent judicial-like entities provides an important counterforce to the power of the community to change the system’s basic standards through externally-initiated deliberative processes. The algorithmic encoding of such external dependence raises various challenges, which have been discussed in the literature on human-centered artificial intelligence. They include, for example, (a) designing a robust and verifiable communication channel between the AI agent and the external entity; (b) allowing informative feedback between the AI agent and the external entity (which would, for example, enable the AI agent to ask for clarifications when the external entity’s decision is ambiguous); and (c) constructing default rules that would determine what the AI agent should do if the input it receives does not provide an answer to the question it posed or in cases in which the external entity fails to provide an input.

121. See Endicott, supra note 26, at 33 (noting the important role of judicial institutions in reconciling incommensurable considerations).
The second option through which deliberation can be incorporated into the metalevel code is based on allowing it to directly engage with its users, relying on various deliberative procedures. Such mechanisms should allow the AI agent to resolve dilemmas by directly approaching the users without recourse to intermediary institutions. Such procedures can be initiated by the AI agent or the public. In the former case, the AI agent can rely, for example, on deliberative polling, in which a representative sample of the relevant population is recruited using stratified random sampling and is then convened to deliberate and decide on a chosen problem. In the latter case, the system can incorporate a petition system in which the AI agent must initiate a debate on the topic once a certain number of people join a petition. Crowdsourcing techniques are already used by other social media platforms, such as Reddit.

A key difficulty in the foregoing framework concerns the power of the AI agent to oversee the deliberation process. According to our model, the AI agent is algorithmically bound by the results of the deliberation process, but it could influence the outcome of that process by intervening in its design and dynamic. The extent of such intervention depends on the autonomy of the meta-agent. Recall that we conceptualized the metalevel algorithm as a political satisficer, which aims to satisfice human preferences, subject to deliberative constraints. The idea of political satisficing can be interpreted as allowing the AI agent to intervene in the deliberation process if this intervention generates a better understanding of human preferences. Such intervention can range from choosing the sampling technique that is used in deliberative polling processes to determining the scale and scope of the information provided to the people participating in the deliberation or to the panelists of an external out-of-court dispute settlement body (taking into account humans’ limited capacity to read


significant amounts of data).\textsuperscript{126} Thus, trying to bind the AI agent to social deliberation paradoxically gives it partial control over the very process by which it is supposed to be bound. It is at this junction that societal and algorithmic constitutionalism collide. The challenge we face can therefore be articulated as follows: Can we design a legitimate, fair, and safe constitutional law system cogoverned by humans and AI algorithms?

V. CONCLUSION

The idea of algorithmic constitutionalism, which we developed in this article, offers a response to the risks of algorithmic governance. It is based on three pillars: (a) layered architecture consisting of object-(operative) and metalevels, (b) algorithmic metareasoning, and (c) correction by deliberation. Our approach is based on the thesis that the AI governance system should be infused with uncertainty about human preferences, which should consequently be translated into uncertainty about the correct interpretation of system rules and overarching mission. We argue that the system should have a variety of hardwired procedures through which it may receive input from its users about the application of the standards, the fundamental principles of the system, and the potential revisions of these rules. But, as we elaborated above, paradoxically, the attempt to subject the AI algorithm to external deliberative control also opens the door for the AI agent to intervene in that process, potentially undermining its very purpose. This outcome reflects a tension between societal and algorithmic constitutionalism, which we believe poses a critical new challenge to constitutional law.

Our article constitutes only a first step in articulating the idea of algorithmic constitutionalism, which raises further legal, philosophical, and technological questions. We hope it will pave the way for further investigations of this concept. A key challenge is to examine how the theoretical concept of algorithmic constitutionalism can be implemented in practice. One possible direction for research in this context concerns the idea of modular implementation. Is it possible to limit the application of algorithmic constitutionalism to those aspects of the platform’s governance, which are more critical from a human rights perspective (e.g., content

\textsuperscript{126} Meta Oversight Board’s decision on the cross-check system illustrates this risk. The Board criticized Meta for failing to provide it with sufficient details on the entities that receive privileged protection under the ERSR system. See OVERSIGHT BOARD, POLICY ADVISORY OPINION ON META’S CROSS-CHECK PROGRAM 27 (2022), https://www.oversightboard.com/news/501654971916288-oversight-board-publishes-policy-advisory-opinion-on-meta-s-cross-check-program/.
moderation), and give the controlling corporation (e.g., Meta) more leeway to intervene in the regulation of less sensitive aspects (e.g., ad delivery)? Such modular implementation could provide a way to balance between the corporation’s legitimate business interests and its human rights commitments and, consequently, could make profit-seeking corporations, such as Meta, more receptive to the idea of algorithmic constitutionalism.¹²⁷

Another potential direction for research is to examine the relations between algorithmic constitutionalism and emerging regulatory initiatives focusing on the risks of AI. Particularly relevant to this inquiry is the new DSA, which European Parliament approved in October 2022.¹²⁸ The DSA aims to extend the responsibilities and obligations of providers of information services, including digital platforms, such as Facebook (defined as “hosting services” in the DSA).¹²⁹ One of the key contributions of the DSA is to establish a new regulatory framework for content moderation. This framework creates a semiconstitutional structure that defines the metaprocedural framework for governing digital platform content moderation actions. The structure includes, for example, provisions regarding the notification and provision of reasons (regarding, for example, the removal of content, disabling access to content, or demoting content),¹³⁰ the establishment of an internal complaint-handling system that enables users to lodge complaints electronically and free of charge.

¹²⁷ A sharper distinction between the corporation’s business and human rights commitments could also be a source of extra costs. For example, the Oversight Board recommended in its cross-check opinion that Meta should prioritize expression important for human rights and make sure that the review of such content does not compete with Meta’s business partners for limited resources. This recommendation requires Meta either to expand its review capacity or to discontinue its current practice of giving special privileges to users with many followers (and are, therefore, a significant source of revenue). See id at 5, 27, 37.


¹²⁹ Digital Services Act, supra note 11, art. 3(g). For a general discussion of the DSA see, Jorge Morais Carvalho et al., Introduction to the Digital Services Act, Content Moderation and Consumer Protection, 3 REVISTA DE DIREITO E TECNOLOGIA 71, 89 (2021); A. R. Lodder & J. Morais Carvalho, Online Platforms: Towards an Information Tsunami with New Requirements on Moderation, Ranking, and Traceability 537, 33 EUR. BUS. L. REV. 537, 547 (2022).

¹³⁰ Digital Services Act, supra note 11, art. 17(1)(a).
against a decision taken by the provider,\textsuperscript{131} and the creation of an out-of-court dispute settlement system that would provide users with accessible redress.\textsuperscript{132} The drafters of the DSA have made a significant effort to distinguish between automated and human decision-making. For example, the DSA requires providers of online platforms to ensure that the decisions made by the internal complaint-handling system “are taken under the supervision of appropriately qualified staff, and not solely on the basis of automated means.” It also requires those providers to ensure that recipients of the service can communicate directly and rapidly with the provider in a user-friendly manner, which does not rely solely on automated tools.\textsuperscript{133} Providers are also required to include information on the use of automated means in the notifications they send to users.\textsuperscript{134}

The new regulatory system created by the DSA brings to the fore the tension between algorithmic and human decision-making, which has so far remained outside the reach of formal regulation, allowing digital platforms, such as Facebook, to be opaque about the exact role of algorithms in their content-moderation decisions. The DSA provides a unique opportunity for exploring the tension between algorithmic and societal constitutionalism in a real socio-technical environment.

\textsuperscript{131} Id. art. 20(1).
\textsuperscript{132} Id. art. 21.
\textsuperscript{133} Id. art. 12(1), 20(6).
\textsuperscript{134} Id. art. 16(6).
Political Autonomy in the Digital World: 
From Data Ownership to Digital
Constitutionalism

DAN WIELSCH

Within the course of digitalisation, data have become a substantial factor of production. They serve not only as a resource for the creation of surplus value but also change the production process itself. Through self-learning algorithms it is even possible to innovate production processes without human intervention. Thus, data are taking their place alongside labour and capital. This relevance of data for creating surplus value—both upstream and downstream—on the one hand, and the (technological) possibilities to appropriate them de facto on the other hand, have triggered a discussion about the need and the legitimacy for creating exclusive rights to data or “data ownership.”

However, legal theorising about data property rights would not just have to develop an adequate legal concept of “data” and explicate the function of data property rights for individual and social autonomies but also align a possible data ownership with the broader idea of social ordering through property rights. If the law takes the social function of rights seriously, it must consider the normativity of social orders, which are constituted through the exercise of rights, and, in addition, it must make sure that those who are affected by these orders can participate in shaping them. As soon as the question regarding the function of individual rights for social practice is raised, the question of the implications of the normativity of this practice for the rights (of third parties in particular) also arises; correspondingly, the question about the participation of these rights holders in the formation of the rules for social practice arises as well. Rights enable both the constitution and the transformation of social order. To the extent private rights would allow for changing the rules of social order, private rights act as political rights.
In the area of property law, jurisprudence has a particular tendency toward reification. Apart from the inclination to conceive of freedoms as spatial spheres of dominion, the authority of exclusive rights attach objects as a matter of course. Property rights are regarded as rights that assign the authority to use goods.¹ In so far as data are classified as goods, it only remains for the law to consider which regime of exclusive rights they must be placed under—general property law or intellectual property law, which share the assumption that data should be managed as a personal asset. In consequence, the discussion about data ownership is narrowed to economic participation in the use of data.

A. Data as Information

At first glance, it seems the representation of data as “information encoded in a way that can be processed by machines” attains adequate complexity for legal conceptualisation.² Such an account of data is based on a reification of information designed for the purposes of law. It adopts the concepts of information developed in other sciences and uses the distinction between various types of information for normative statements about the extent to which a certain piece of information should be made the object of exclusive rights. Accordingly, a distinction can be made between semantic information, which is delimited by its meaning; syntactic information, which is delimited through its representation as a set of signs; and structural information, which is delimited by its material embodiment.³ These delimitations are inspired by the distinction between signs and meaning in semiotics, further supplemented by a category of information independent from consciousness but contained in a physical system. These three perspectives on information are not mutually exclusive but stand in a tiered relationship to one another insofar as structural information can serve as a kind of carrier of syntactical information, and this, in turn, as a carrier of semantic information.⁴

With such an “opening” of the concept of information into multiple dimensions, it would be theoretically possible to determine the characteristics of the sociotechnological subject area of data property rights and to give a correspondingly differentiated and substantial

¹. Armen Alchian, Some Economics of Property Rights, 30 II POLITICO 816, 818 (1965).
³. Id. at 194.
⁴. Id.
reconstruction of their effects on social processes. As is recognized, the multidimensionality of the suggested concept of information is like the “layer model” commonly used in the architecture of computer networks, whereby the “physical layer” corresponds to the structural information, the “code layer” to the syntactical information, and the “content layer” to the semantic information.5

However, like the bundle-of-rights model of property, the layered model of communication networks is not simply an analytical tool. Its use by lawyers makes visible the normative scope in the design of network architecture and draws attention to the fact that the question of management through property rights or through commons arises anew at each level. Thus, in the early phase of internet development, the consistent implementation of the “end-to-end” principle in the construction of network protocols6 decentralised the decision-making competence over the use of the “code layer” so that the logical layer is accordingly constituted as a common good.7 Users only have to comply with the specifications of the IP protocol and not with the will of a network operator and the technical specifications required by that operator.8 Such a separation of property (on the physical layer of the network infrastructure) and control (over the possibilities of action on the logical layer) ensures the medium’s openness to development and an enormous variety of network-based applications. With the help of the layered model, it is possible to illustrate how technologically implemented access rules create a new “space” of possibilities of action that are independent of the ownership of, and thus also of the determining powers over, the physical network. A communication medium emerges that can develop autonomously insofar as the power over the material carriers of data transport does not “leverage” to the logical, syntactical level. Thus, in jurisprudence the layered model makes it possible to describe the constitution of media freedoms of use in digital networks as well as possible threats for them.9

The representation of data as goods cannot incorporate the relations that exist between the different operative contexts of data. In its quest for objective delimitation, the data-as-goods model ignores the social

5. Id.
8. See Van Schewick, supra note 6, at 204 (explaining the increase of the independence of developers and enabling autonomous innovation under a broad end-to-end principle).
meaning of data. Assuming the use value of data, the property-law approach focuses on the act of data creation through the coding of information into signs. In line with the model of intellectual property law, which grants the creator as the translator of meaning into signs an exclusive right to the syntactical work created (e.g., an author who writes a story down as text), the person responsible for the coding of information from structural to syntactical should also deserve an exclusive right to the syntactical information. This normative assignment of the positive externalities of data use would replace the current state of de facto exclusivity and promote the trade in data on the basis of a legally secure primary allocation.10 Thus, the potential associated with data for shaping the environment is concentrated in the hands of the data producer. Since the creation of data, unlike works under copyright law, is not bound to specific semantic contexts, the power to control the environment is not specifically restricted. This is in sharp contrast to the situation in IP law, where the tailoring of the scope of rights is supposed to serve the flourishing of exactly those intellectual and cultural domains, which form the basis for its creation in the first place.11 Consequently, data-producer rights lead to an unspecific and—taking into consideration the regulatory potential of Big Data—at the same time invasive restriction of the general freedom of action. For this reason, the exclusive rights of data producers to syntactic information are viewed critically not only from a competition law perspective, as they could interfere with equal access of competitors to the content of the data,12 but also with regard to the lack of possibilities for the data subjects who contribute to the generation of data to have a say in the way they are processed.13

11. On this systemic nexus see WIELSCH, supra note 7, at 79.
13. GERMAN DATA ETHICS COMMISSION, OPINION OF THE DATA ETHICS COMMISSION 85 (2019). Here, the complex character of processes of data generation (data production, data enrichment, and data refinement) is correctly pointed out. Relevant contributions by an actor—in particular by natural persons—are supposed to consist, among other things, in the fact that the information stored in the data relates in its meaning to this actor or to an object connected to him, the data was generated by an activity of the actor, or the data was generated by technology which this actor created or in which he invested.
A reconnection of data to their social dimension becomes possible, however, when they are understood not only as binary code but as “cultural facts.” As digitalisation resolves into distinct pieces of information, all behaviour by citizens (private or public) is likely to result in a “total digitalisation of the civic lifeworld.”\textsuperscript{14} All meaningful action occurring in bilateral relations, organisations, and social institutions can be digitized and is therefore affected by data analysis conducted with the aim to reorientate social action on the basis of detected behaviouristic patterns. Given this totality of a digital recreation of the world, a unilateral legal reconstruction of data in the light of just the economic perspective must appear particularly problematic.

Although the representation of data as cultural facts does not overcome law’s reification bias, it rightly shifts the emphasis to data as building blocks of a new digitized world that will deeply impact everyday life and become its real foundation. As a result of the social pervasiveness of regulation based on data, the individual is affected as a citizen and not as a mere market participant. Nevertheless, it is not those affected who are able to shape the new basis of life, but companies and organisations that appropriate the personal data of individuals as a matter of fact. Due to the factual-first appropriation of data, a kind of state of nature has prevailed so far in the shaping of the digital world, resulting in the alienation of citizens who can no longer participate in the shaping of civil society.

In order to change this state of affairs, the individual must be empowered to have a voice in the digital world’s rulemaking. The individual must be entitled to codetermine how the data that he or she has generated in the first place are used. Some scholars, therefore, propose that individuals be legally granted a sui generis intellectual property right to the data generated by their behaviour.\textsuperscript{15} Such an exclusive right would put the citizen in a position to negotiate with companies on the “network norms” that set standards for the use of behavioural data and the companies’ private data regimes. This provides the remarkable impetus to conceptualise data ownership not simply as a property right, but as a “civil right”—much like Sinzheimer’s “economic citizen” should have had the dual right to the


\textsuperscript{15} \textit{Id. at} 360.
freedom of self-determination and political intervention in economic life.\textsuperscript{16} Ownership rights to data may therefore indeed open up the possibility of economic participation,\textsuperscript{17} but first and foremost they can be read as an expression of the political autonomy of the individual to effectively influence what happens with networked data. Since the conditions for individual autonomy can be subject to regulation without necessarily making reference to identifiable natural persons but can also be based on data generated by a person’s use of items in his or her environment, the object of a sui generis data property right is, in contrast to data protection law, to be consequently extended to nonpersonal data.\textsuperscript{18} However, the limitations of data protection are also exceeded insofar as a data property right would have to be obeyed by third parties as well, could be asserted against any unauthorised user in the value chain, and would still apply even if the data which were originally related to a person are later anonymised.

\textbf{C. Data as Medium}

However, the all-encompassing lifeworld conception of data cannot capture the autonomy of the digital medium. Eventually, this conception remains trapped in the premise of action theory that the individual can determine the communication of his behaviour in a sovereign way.\textsuperscript{19} Accordingly, the construction of meaning remains bound to individual acts of understanding and can only register the sociotechnological generation of meaning from its effects but cannot explain its genesis and dynamics. Understanding data as a (technological) “medium” instead pushes further in that direction and radically separates data from the signs they transmit. Data then are not conceived of as information but as a (binary) representation of information. The digital syntax alone allows them to be a means of processing, storing, and transmitting information and to be the object of technological operations.\textsuperscript{20} From this


\textsuperscript{17} Fezer, \textit{supra} note 14, at 364 (providing the lawful use of conduct-generated data requires either an individual licensing agreement, or a form of collective rights management, or a statutory liability rule).

\textsuperscript{18} \textit{Id.} at 360, 364; \textit{Contra} Trakman ET AL., \textit{supra} note 10, at 937 ( focusing only on personal data); For a criticism of new IP rights to non-personal data, see WOLFGANG KERBER, A NEW (INTELLECTUAL) PROPERTY RIGHT FOR NON-PERSONAL DATA? AN ECONOMIC ANALYSIS 98 (2016).

\textsuperscript{19} Fezer, \textit{supra} note 14, at 360, 363.

\textsuperscript{20} Marc Amstutz, \textit{Dateneigentum: Funktion und Form, in ARCHIV FUR DIE CIVILISTISCHE PRAXIS} 438, 452 (2018).
perspective, the semiotic model, which focuses on the content abstracted by data and its usefulness, must arrive at an erroneous analogy with intangible goods. Due to the characteristic of data as public goods, it deduces the establishment of exclusive rights as a prerequisite for markets in data. If, on the other hand, data are regarded as a medium, they would be indispensable for communication in their function of transmitting information, and their production would not be a question of managing a resource but an expression of communication. However, this take on data builds on a particular concept of what a medium is. Following Kittler, one cannot refer to media as signs and symbols but must strictly distinguish between symbolic and technological media, to which data are said to belong.\footnote{Id. at 462.}

The operation of data as a medium also explains their world-creating, socially generative function.\footnote{See Salomé Viljoen, A Relational Theory of Data, 131 YALE L.J. 573, 638 (2021) (The population-based effects of data processing in the digital economy are also accounted for by an approach that conceives of data as social relations and further claims that data’s relationality should be negotiated and managed democratically, thus conceptualizing of “data as democratic medium”).} The reality that data create is, of course, not a representation of social reality.\footnote{Id. at 612; see also Amstutz, supra note 20, at 511 (“Digital reality arises from striving to ‘de-socialise’ social reality in order to reconstruct it statistically with biopolitical intent”).} On the basis of bulk collection of behavioural data, algorithms determine correlations between people and events that serve to make predictions about the future behaviour of people. The accuracy of the predictions is less important than the preemption of possible actions that can guide the further operation of social systems. Autonomous digital systems undermine the function of understanding.\footnote{Amstutz, supra note 20, at 511 (attributing the actions to single individuals for the sake of communication is no longer possible since behavioural data get aggregated into clusters; however, it is no longer necessary either since social life is no longer a matter of interpreting and understanding actions, but of anticipating possibilities for action and technically implementing them.)} Because the preempted possibilities of action are technically enforced, the algorithmically constructed digital world validates itself. In this way, the models of social systems about persons begin to form on the basis of algorithmic rationality. This paves the way for an “algorithmic governmentality” in which the control of behaviour does not take place according to social norms that express a certain expectation and, at best, would have been filtered in deliberative processes, but through technological architectures that anticipate all possibilities for action and are derived
from inductive computational processes.\textsuperscript{25}

Accounts of data as a medium are set out to take into account this “algorithmic governmentality” developing via data. Some proponents try to achieve this through the introduction of data ownership. Like the conceptualization of data as cultural facts, they see the function of a data property right as reappropriating and safeguarding civil autonomies instead of economic benefits arising on the side of corporate data producers. Beyond the protection of personal autonomies, however, the law should also consider the relation between the digital medium and social autonomies. This relationship is to be dogmatically implemented through a data ownership right, which is to be conceived according to the classical model of tangible property.\textsuperscript{26} To be sure, when measured against the basic idea of understanding data as a medium, a recourse to the institute of property has something paradoxical about it. For the consequence of such a qualification would actually be to keep the use of a medium free of individual authority. However, the de facto control of the syntactic level by data gatekeepers provides them with normative influence downstream on the sociosemantic level. To nevertheless continue to guarantee the conditions for free action on this dependent level, the concept of data as media sees itself forced to regulate the syntactic level from there upstream, as it were, and against its own premises.

II. PARTICIPATION RIGHTS IN DIGITAL RULEMAKING

Advanced proposals for the creation of “data ownership” as the ones discussed, which are responsive to the full spectrum of the social dimensions of data, can certainly provide the building blocks for a legal constitution of the digital medium. They reconstruct the concept of data from a nonreductionist, multilateral perspective in law. On this basis, it is possible to consider the effects of autonomous digital systems operating in the medium of data (and not of meaning) for personal and social autonomies alike.

The strategy of seeking to replace the de facto appropriation of data and the associated digital rulemaking power of data companies with de

\textsuperscript{25} Antoinette Rouvroy, \textit{Governing Without Norms: Algorithmic Governmentality}, 32 PSYCHOANALYTICAL NOTEBOOKS 99, 99 (2018) (describing the distinctive nature of this form) (“The ‘profiles/scores/matches’ through which individuals are classed, assessed, recomposed or sanctioned, and through which the merits, needs, opportunities or dangers contained in life forms are assessed, are based on the numerical signals arising from trajectories, relations and interactions of individuals rather than on norms resulting from prior deliberative processes.”)

\textsuperscript{26} Amstutz, \textit{supra} note 20, at 541.
jure data property rights assigned to citizens brings the political dimension of the legal institution of property to the fore. However, in contrast to the conventional use of property rights within the framework of a market constitution, the function of data ownership would be to enable a legal reconstitution of social practice. In this respect, the concept of data ownership is more akin to the reorientation of property rights through open source licenses. The crucial point here, as there, is that these rights provide the instruments for the modification of a social practice that can be set in motion by individuals. This modification is a matter of “counter-rights” to enable an alternative social practice.27 The exclusion right serves less to internalise the positive external effects of use and rather to help shape “network norms,” which decide in particular how social models of the person are generated.28 In the foreground of the digital participation rights that the mentioned approaches refer to as data ownership is the participation in the legal constitution of the digital medium, which is effective throughout society, and the central task of which is the protection of the “person.” In this respect, these rights seem, if anything, to be functionally equivalent to data protection rights.

In the eyes of the advocates of data ownership, however, data protection law, at least in its present form, is not up to the task of regulating digitalisation in civil society.29 This is because consent under data protection law—and a “data contract law” based on this—cannot function as a source of legitimisation for mass data collection with regulatory effects for third parties nor for cross-contextual data networking. This networking can lead to the transgression of context-related information norms and thus to the violation of the integrity of social contexts.30 The weak spot of the General Data Protection Regulation (GDPR) consists in the strategic and statistical third-party effects of individual consent to data processing.31 However, while a data ownership law promises to provide the data subject with more effective

28. How to grant people a say in the social processes of their own formation is also the reorientation of data governance pursued by the account of data as democratic medium. See Viljoen, supra note 22, at 641.
30. On “transgression” of information norms see HELEN NISSENBAUM, PRIVACY IN CONTEXT 186 (Stanford University Press 2010).
control than data protection law in some situations,\textsuperscript{32} it shares the problems of consent in data protection law as a result of its individualistic conception. Considerable information asymmetries between the contracting parties, disclosure dynamics through the behaviour of third parties according to rational choice theory, and absence of protection-related competition on the merits in data-intensive markets also affect the validity of consent in transactions about data property. As a result of the deficits of isolated consent requirements on the one hand and collective action problems on the other, the data ownership model is forced to set up a representative body for the collective exercise of rights in order to guarantee effective data control.\textsuperscript{33}

Eventually, as in the case of legal consent under data protection law, substantive limits would have to be set on the individual data owner’s power of disposition, and, just as in that case, fundamental rights provide decisive normative guidance for these limits. Contractual consent to the use of data as well as the reasonableness of standard terms for digital services must be strictly interpreted in conformity with fundamental rights. Hereby, of course, neither a formal-liberal, a welfare-state material, nor a functional, institutional interpretation of basic rights is to be assumed, but rather a relational one, as is the case with other individual rights.\textsuperscript{34} Fundamental rights require private rulemaking to become reflexive through procedural requirements that oblige these rule-makers as well to consider all autonomies affected by the stipulated rules.\textsuperscript{35}

If one takes into account that all rights must be tied back to the social context in which they are exercised, then their concrete dogmatic instantiation is ultimately of secondary importance for the concern of civil society to participate in the shaping of society. In the end, there could be a bundle of effective digital participation rights that run across different areas of law. The concept of ownership makes a point in this respect because it places an emphasis on the competence to shape one’s own environment. However, the task of a digital constitutionalism must

\textsuperscript{32} Patrik Hummel et al., \textit{Own Data? Ethical Reflections on Data Ownership}, 34 PHIL. & TECH. 545, 553 (2021); Fezer, supra note 14, at 366.

\textsuperscript{33} But see Karl-Heinz Fezer, \textit{Digitales Dateneigentum in DATENEGENTUM UND DATENHANDEL} 101, 152 (Stiftung Datenschutz ed., 2019). According to Fezer, the exercise of data ownership rights by a representative body is not a matter of an instrument of collective legal protection. Because he understands data ownership not merely as a right of market participants, but as a civil right, he aligns himself with the democratic principle of representation for the exercise of creative competence.

\textsuperscript{34} For a relational theory of individual rights, see Dan Wielsch, \textit{Relational Justice}, 76 L. & CONTEMP. PROBS. 191, 191 (2013).

\textsuperscript{35} Id. at 198.
not be identified with a specific legal institution and its doctrine must not be identified with property intangibles nor with intellectual property. The concrete recourse to the property form for the control of digital regulation appears problematic because individual data have no use value for which appropriation possibility should be protected. Moreover, the Lockean idea that the labour invested in the creation of things gives rise to property rights, which is effective in both legal fields, has little persuasive power in view of the automated aggregation and evaluation of data. Above all, however, the application of the property form tempts us to regard data as an object and, in so doing, to detach data from the social practice in which they arise and have effects. The property form would itself bring about just that reification of social regulatory data practice, which the law must reflexively abolish for the sake of the autonomy of the data subjects. In order to avoid this narrowing by the property form and, conversely, to avoid a dysfunctional adaptation of property doctrine to data as legal objects, civil rights of participation should be directly related to the formation of rules for the digital medium. This formation should be in accordance with the idea of social constitutions and dogmatised as “digital rights” specific to each subject area. What is decisive is the development of a level of reflection in law that can take into account the autonomy of the digital medium and place it in relation to other autonomies.

Because the digital syntactic world influences social differentiation and thus in turn the preconditions of individual freedom, the digital world acquires its own normative quality. The independence of this new digital normativity from the analogue world must not lead to its naturalisation and be regarded as a given. The new digital normativity forms in the first place as the subject matter of a justification process from the perspective of society as a whole. Appropriately set digital rights and obligations of data subjects and data processors act irrespective of their dogmatic instantiation as part of a media constitutional law, if their function is to constitute the autonomy of the digital medium, while at the same time limiting it in relation to other autonomies. This is similar to how the creation of society’s political self-determination in the state constitution is at the same time linked to a protection of social autonomies from the political. Only from this perspective of digital constitutionalism does it become obvious why legal

36. Dan Wielsch, Die Ordnungen der Netzwerke in 80 DIE ZUKUNFT DES NETZDG UND SEINE FOLGEN FÜR DIE NETZWERKKOMMUNIKATION 61, 63 (Martin Eifert & Tobias Gostomzyk eds., 2018).

37. On the constitutive and limitative functions of societal constitutions see GUNThER TEBuNER, CONSTITUTIONAL FRAGMENTS: SOCIETAL CONSTITUTIONALISM AND GLOBALIZATION 75–88 (Gareth Norbury trans., 2012).
principles such as data transparency, data security, data interoperability, and data transportability, which are already stated in data protection law, are said to also form part of the data property proposed. They are not limited to the field of data protection law but represent principles of digital constitutionalism that apply across the different legal areas. They form the subject matter of data-related codetermination rights, which are also to be made effective in cases of digital regulation. In this respect, one could speak of a “civil society effet utile” of digital rights because the new control dispositive has a primarily societal character. It is in the hands of private operators of algorithmic systems—insofar as control has not already been handed over to self-learning algorithms—and precisely not in the hands of the state.

Under current law, important rulemaking participation rights of data subjects are already established for personal data among others in the form of the detailed consent requirements making sure that it is freely given (Art. 6 (1)(a), Art. 7 (4) GDPR), in the form of the possibility to withdraw consent at any time (Art. 7(3) GDPR), or in the form of the right to have personal data erased (Art. 17(1) GDPR), being made effective even vis-à-vis third parties by means of a right to forward the erasure request (Art. 17(2) GDPR). Also in the case of nonpersonal data, a de facto regulatory power of data-controlling gatekeepers is normatively broken up by competition-specific access rights as stated in Art. 6(8), (10), (11) of the Digital Markets Act (DMA) and § 19(2) No. 4, § 20(1)(a) in conjunction with § 20(1) and § 19(a)(2) No. 5 of the German Act Against Restraints of Competition (GWB). Beyond gatekeeper constellations, it is generally proposed to assume a special legal relationship between actors with factual data control and those generating the data in the first place (such as the owners of sensor-equipped devices), from which rights of codetermination can be derived. Even the third-party effects of contractual agreements on data use could be regulated in the future, for instance along the lines of Art. 4(4) of the Trade Secrets Directive (EU) 2016/943. In contrast, granting propertarian entitlement to data would, from a data protection perspective, create problematic financial incentives for the generation of

39. 2022 O.J. (L 265) 1.
40. GERMAN DATA ETHICS COMMISSION, supra note 13, 147 (Factors giving rise to data rights may consist in: extent and nature of the contribution to data generation; weight of the individual interest in the granting of the data right; weight of conflicting individual interests of the obligated party; weight of general interests; distribution of power between the participants in the special connection).
data, especially by individuals, and completely link the processes of data use to the dynamics of the market economy.\textsuperscript{41}

To be sure, the creation of digital participation rights for data subjects goes hand in hand with an increase in transaction costs. This increased cost is not only to be accepted, but is, in a way, part of the idea behind constitutionalizing digital media. The proposal is to create negotiation mandates which digital services would have to take into account in the construction of their privately crafted data regimes, even if, and precisely because, the participation of those affected by the regimes could result in different designs than in the case of usual unilateral specification by service providers. In a parallel with data protection law, one can speak of an “intended market failure,”\textsuperscript{42} in which precisely the extended options for codetermination are reflected. This is neglected by those who argue that data ownership would “entail enormous transaction costs without any recognisable macroeconomic benefit in return.”\textsuperscript{43} Such a view narrows the discourse on the justification of digital rights to a purely economic rationale and thus misses the level of reflection required of law by society as a whole.

The economic argument against data ownership is, of course, on its own merits conclusive in its focus just on the particular social institution of functioning data markets. According to this argument, de facto possibilities of appropriating data already allow contractual compensation for the costs of data processing so that no property right is needed to establish functioning markets. On the contrary, the legal protection of raw data necessarily has the effect of restricting access to and the use of data even at the semantic level, where the actual creation of added value in the form of the creation of new knowledge takes place.\textsuperscript{44} It is indeed precisely the enabling of such a “leveraging” of control from the syntactic to the (sociogenerative) semantic level that is at stake in the participation rights under data constitutional law.

Faced with the de facto appropriation possibilities of the data producers, the economic approach also sees itself forced to postulate participation rights in remarkable parallel, even though tailored to just one particular institution. A “data access right” is proposed for those persons from whom data is collected and “who therefore have a special

\textsuperscript{41} Cf. id. at 104 (against data ownership but explicitly in favor of co-determining and participatory rights).
\textsuperscript{43} Josef Drexl, \textit{Neue Regeln für die Europäische Datenwirtschaft?}, 2 NZKARTR 339, 341 (2017).
\textsuperscript{44} Id. at 343.
interest in access to this data for the purpose of further data evaluation.” The purpose of such digital rights, doctrinally constructed as access rights, is to ensure a free flow of data in the sense of “data sharing,” which also benefits competitors in downstream data markets. Thus, the idea of data access rights is—in the spirit of a digital constitutionalism—extended in two ways: on the one hand, it is not to be linked to a market-dominating position of the company collecting the data and, on the other hand, such a data access right proves to be a generalisation of the right to data portability in Article 20 of the GDPR, which is thus carried beyond the scope of the GDPR into other areas of regulation such as competition policy.

III. STATUS ACTIVUS IN THE DIGITAL WORLD

Digital conditions reveal the grandiosity of private law and at the same time the danger of its continued tragedy. It is able to penetrate social fields that are closed to sovereign administration, but in order to do so it must fully recognise its political dimension and not see itself merely as market regulation law. It was in private law that the promise of sociopolitical emancipation once rested. The political postulate that everyone in society should have the same rights and the same status could—as was assumed—be fulfilled by affording each individual the status of a person under private law. However, the requirements for the success of this idea were immediately buried again in the legal thinking of the nineteenth century, insofar as it shifted subjectivity, with Kant, into two worlds in which “freedom (as the freedom of man’s autonomous morality) and citizen (as ‘co-legislator’ in political society) move apart.” Appropriate legal instruments are needed to bring these worlds together again by understanding private rights as rights to participate in the formation of social rules and to release the political autonomy associated with them. The legal matters that operate as parts of a digital constitution must confer a suitable status activus on the data

45. Id. at 344.
46. Id.
47. Franz Böhm, Privatrechtsgesellschaft und Marktwirtschaft, in 17 ORDO 75, 75–76, 78, 80 (1966) (There should be “in the sphere of society only one legal status, equal for all, only one competence of planning and acting, equal for all, namely private autonomy.”).
48. Rudolf Wiethölter, Bürgerliches Recht, in Handlexikon zur Rechtswissenschaft 50, 53 (Görlitz ed., 1972) (translation by author). (This corresponds to the historical development: “The replacement of the absolutist-feudalist society did not lead to a political civil society with corresponding civil law, but to the formal-liberal civil constitutional state, i.e. to the dualism of a ‘political’ authoritarian state and an ‘apolitical’ laissez-faire society determined by the propertied and educated bourgeoisie.”).
The special authority associated with property to shape the environment hinges on the exclusive allocation of comprehensive rule-setting powers in relation to an object. Resulting both from the social effects of use by the owner and from the dependency of nonowners on use, a considerable social-regulatory potential arises from the normative allocation of exclusive decision-making powers—both factors gaining particular social relevance in the case of property in corporations. This legal ontological framework has become fragile in two ways due to the digitalisation of the world and the emergence of a data economy. On the one hand, the regulatory power of data as a factor of production is not based on the ownership of finite goods anymore but on the translation of social relations into the medium of data. It is not individual—corporeal or incorporeal—objects that are appropriated but social interaction and practice. On the other hand, unlike object-based production, the “commodification of behaviour” through the data economy does not require the “ownership of objectified labour.” Data-driven regulatory power does not require ownership of the data as a factor of production. For data companies, de facto control is sufficient to enable decisions about the purposes and means of processing data. The codetermination opportunities of digital rights are to be adjusted to this by removing, changing, or sharing such technology-based control. Unlike codetermination rights under labour constitution law, they are not Dioscuri of property. However, they share its basic concept of constitutionalizing an autonomy the operation of which has broad transindividual effects and essentially determines the conditions for a self-determined life of others: the autonomy of the corporate actor and the autonomy of the digital medium respectively.

Since the time of the bourgeois revolution, the category of property has refused any intermediation of freedom and instead patronized the indivisibility of the individual will. Insofar as legal scholarship draws on it to deal with digital regulation, it would have to transform itself beyond recognition, at least according to continental understanding, by dissolving the unitary form and distributing various rights of control among the different stakeholders of data processing. As it turns out,

49. Cf. Goerg Jellinek, System der subjektiven öffentlichen Rechte 132 (1892) (on status activus in relation to the state).
51. Formulation based on GDPR Art. 4 No. 7 (2016) for the identification of “controller” in the English language version.
52. Sebastian Lohsse et al., Consistently van Erp, Management as Ownership of Data, in Data as Counter-Performance – Contract Law 2.0? 77, 91 (2020) (according to which the object qualifies the law and not vice versa) (The claims, which are to be
the concept of data ownership is neither required for normative reasons nor is it dogmatically appropriate, neither on the side of the producers nor the data subjects. For it is not the rights that qualify the medium, but the medium that qualifies the rights. Codetermination through digital rights would realise the autonomy of data citizens in a world without property.

associated with a “fluid” new concept of property, consist of “access, control (exclusion), portability and erasure” and authorise the “management” of data.)
Rage Against the Machine: Profiling and Power in the Data Economy

IRINA DOMURATH*

ABSTRACT

This contribution deals with algorithmic profiling as an example of datafication and machine colonization. It examines to what extent Teubner's theory of societal constitutionalism (SC) describes the emergence of an EU digital constitution, including its capacities to deal with datafication and machine colonization. This article advances an internal critique of SC, by questioning its assumptions concerning the capacity of societal actors and nonlegal media, such as public outrage and litigation, to exert the pressure needed for change from within. Using Habermas’s colonization theory as well as the insights of the law and political economy literature helps to understand the structural power of companies as an inhibition to the buildup of external pressure and to justify the adoption of a counter-concept of structural digital vulnerability, which can be used to interpret existing rules in a stricter way.

Keywords: Profiling, Datafication, Machine Colonization, Societal Constitutionalism, Teubner, Habermas, Political Economy

I. INTRODUCTION

The development of the data economy is probably one of the most disruptive events of our time. Social media platforms and digital management practices deeply impact the way in which the economy works, shapes, and reshapes social relations in almost all spheres: finance, commerce, work, housing, family, and friends. Even the

* Adjunct Professor of Law, Universidad Adolfo Ibañez, Chile. I thank Angelo Golia and Gunther Teubner for their invitation to participate in this project and for indispensable comments on earlier drafts. I also thank Poul Kjær, Marta Maroni, Felipe Padilla, Stefano Palestini, and Federico Szczaranski for their insights and helpful discussions of the topic. All errors are mine.
traditional democratic ways of knowledge creation and political information are changing.

The rise of transnational corporations and global governance regimes has led to a distinctive type of norms. The data economy’s governance structures are established by an interplay of national and supranational norms as well as private economic actors’ terms and conditions and privacy policies. The term algorithmic regulation, more specifically, denotes decision-making systems that regulate a domain of activity to manage risk or alter behavior through the continual and computational generation of knowledge from data.¹ It is a subset of regulatory governance. This computational turn in government is understood as the calculational framing of the possible actions of others by both public authorities and private actors.²

Constitutionalism beyond the state is the intellectual movement that has tried to theoretically grasp the legal orders that arise beyond the borders of the nation states. Many try to reread the founding treaties of international organizations as constitutions for new global orders,³ including the UN Charter as a world constitution.⁴ Similarly, principles and human rights are seen as traits of the constitutionalization of transnational legal orders, both public and private.⁵ Teubner’s SC has a specific take on such legal orders. His theory claims that legal regimes established by global private actors have a constitutional character outside of the institutionalized political sector, which leads to a fundamental transformation of preexisting constitutional orders.⁶ SC is different from other theories of transnational constitutionalism because it assumes that the world

society is fragmented into nations, transnational regimes, and regional cultures as well as functionally differentiated communication systems. Here lies SC’s appeal: as an all-encompassing theory, it draws a broad picture and includes—inter alia—the political system (based on the medium of power turning around states) as well as the economic system (based on the medium of money and shaped by both public and private actors) and machine systems (with data as their medium).

SC argues that societal constitutions can emerge in different (sub-)systems. Their goal is to limit the expansionist tendencies of functionally differentiated systems, similar to the limitation of political power through political constitutions. Broadly, it deals with the question of whether and how external pressures can counteract the expansionist—colonizing and instrumentalizing—tendencies of autonomous subsystems. This is possible because it assumes, in line with Luhmann, that systems are cognitively open, which means that they are influenced by other systems through coupling. The constitutional question of SC points at possibilities for internal self-limitation of autonomous subsystems to avoid the colonization of others.

This contribution is a thought experiment. I am interested in testing whether SC can deal with the regulatory challenges of algorithmic governance in the data economy, which I have characterized elsewhere as totalitarian. In the language of systems theory, the data economy is a subsystem of the global economic system. In structural coupling with machine systems, it is driven by datafication, the unrestricted expansion of data as the communicative means of machines, or the “digital capture of social activities for the use, transformation and manipulation of subjects and objects” through the combination with other information infrastructures (for example, the monetary system). To this end, data is continuously collected and combined with metadata. As more and more goods and services are nowadays based on or connected with data collection, the danger of machine systems colonizing other systems ensues. Here, SC comes in, asking how this colonization can be prevented.

The underlying question informing this contribution is whether

---

7. *Id.* at 14.
9. *Id.*
external pressures can counteract the expansionist—colonizing and instrumentalizing—tendencies of machine systems. In other words, to what extent can SC prevent colonization through machines? I hypothesize that SC's strong (albeit not exclusive) reliance on external pressure from societal actors undermines its own potential in the face of unprecedented power asymmetries between actors in the digital sphere. Rage against the machine as external pressure for change might not develop.

In Section II, I will outline the constitutional question in the digital sphere from the perspective of SC. To this end, I first summarize the characteristics of SC (II.A) before making the case for consumer profiling as an example of datafication and machine colonization (II.B). In Section III, I sketch the rules of EU digital law—understood as the body of rules applicable to the data economy generally and profiling specifically—in terms of SC. Section IV outlines an internal critique of SC by questioning its assumptions. In order to advance this critique, I rely, first, on Habermas' colonization theory, which elucidates the problem of expansive machine systems in terms of ensuing pathologies, and second on the insights of law and political economy (LPE), which shed light on how unprecedented corporate power in the digital sphere and datafication undermine the very basis for the normative dimension of SC. In final Section V, I outline an approach to tilting the power in the digital sphere. I argue that SC can work, if a stricter interpretation of existing rules is conceptually underpinned by the concept of structural vulnerability.

II. THE CONSTITUTIONAL QUESTION IN REFLEXIVE DIGITAL LAW

A. What a Societal Constitution Is and Does

SC aims at enabling reflexivity and stimulating processes of communication, which are sustained externally but processed internally (through generalization and respecification). In this way, a functionally understood constitution may emerge that is able to ensure both the autonomy of the system's communicative medium as well as the limitation of its expansive tendencies. Emphasis lies on the necessity to limit the expansive tendencies of systems. To counteract

13. I am limiting my analysis to two of the four transnational constitutional norms of SC, leaving aside constitutional processes of double reflexivity and structures in the forms of hybrid meta codes.
15. TEUBNER, supra note 6, at 76.
their pathological growth, Teubner advocates for external pressure toward self-limitation to achieve learning and self-correction, arguably “at the very last minute.”\textsuperscript{16} Teubner calls the recognition of the destructive growth tendencies the “constitutional moment”—the immediate experience of crisis.\textsuperscript{17} In this crisis, pathologies emerge when disturbances in the communication within and among systems cannot be productively exploited. As a result, irritation among systems and mutual limitation are inhibited and systems continue expanding. This “pathological growth” is a process that goes beyond the need for continuity of the system ("normal growth") but, instead, threatens the maintenance of the system.\textsuperscript{18} For example, the excessive growth of the welfare state is a pathology in the political system. Similarly, the incessant growth of research prompting more research is a pathology in the science system.\textsuperscript{19} It is the system’s pathological tendencies that bring about a crisis in which the decision between destruction and self-restraint must be taken.\textsuperscript{20} The result of this self-reflective process must be a change in the “very arteries of the communication circulation,” a “capillary constitutionalization.”\textsuperscript{21} In the end, SC seeks to establish and thrive on a delicate equilibrium between colonization and autonomy. Teubner calls this a “difficult balancing act.” The constitutional moment is the moment of crisis\textsuperscript{22} in which—“with a bit of luck”\textsuperscript{23}—this kind of synthesis can be achieved.

The question that orientates SC is how external pressures can become strong enough to push subsystems toward self-limitation. Teubner emphasizes that the task of coordinating the function of the system and its environment is difficult and can only be accomplished through internal and experimental self-reflection, and not through external and planned intervention. The latter can only prompt but not replace the former.\textsuperscript{24}

Societal actors and public outrage play a pivotal role in bringing about change. State power must function together with external societal forces to exert such massive pressure on the expansionist, functional system that the system will be forced to limit its expansive tendencies (through self-direction).\textsuperscript{25} This gives the civil society’s forces such as

\textsuperscript{16} Id. at 82.
\textsuperscript{17} Id. at 81–82.
\textsuperscript{18} Id. at 80–81.
\textsuperscript{19} Id. at 81.
\textsuperscript{20} Id. at 82.
\textsuperscript{21} Id. at 83–84.
\textsuperscript{22} Id. at 82–83.
\textsuperscript{23} Id. at 84–85.
\textsuperscript{24} Id. at 84.
\textsuperscript{25} Id.
media, public discourse and outrage, protest movements, NGOs, and the like the important task of acting as a counter-power to expansionist forces. Public outrage against machine colonization acts as the impulse for internal reflection and self-limitation.

This idea rests on the rejection of pure state regulation, which is based on the experience with the totalitarian regimes of the twentieth century. It also rejects the Habermasian view of law as an institution that provides a normative standard and facilitates communicative processes by guaranteeing an external constitution. Law as an institution reaches its goals through law as a medium—that is, norms of procedure and legal protection in order to reach procedural legitimacy. In contrast, Teubner argues that (in the light of functional differentiation) universal legitimation structures (as a general morality of discourse or a common procedure of reflection) are not possible anymore, mainly because the legal prerequisites for reflection processes differ in each system.

As a consequence, SC relies on procedural norms that regulate processes, organization, and distribution of rights and competencies that are to indirectly and abstractly control social action. Law as an autopoietic social system consisting of communication must act at the subsystem-specific level to install, correct, and redefine democratic self-regulatory mechanisms, a type of self-regulation that takes into account the specificities of different communication media of the functionally differentiated systems. The integrative function of contemporary responsive law is to realize its own reflexive orientation by providing the structural premises for reflexive processes in other social subsystems. As such, Teubner’s theory is highly constructivist in orientation: law is seen as an autonomous system that constructs a social reality of its own instead of responding to some other social reality. It is “living law” that oscillates continuously between autonomy and heteronomy.

26. Id. at 83.
28. Id. at 544–45.
30. Id. at 255.
33. Teubner, supra note 33, at 730.
34. Golia & Teubner, supra note 8, at 4; id.
B. Profiling and Datafication as a Case of Application

Profiling is an almost obvious example of an application for SC in the digital sphere. It reflects the algorithmic turn in government and private rulemaking as well as pathological datafication and expansion of machine systems. The constitutional question that SC asks here is to what extent and how machine colonization can be limited.

At its most basic level, profiling distinguishes people into different categories based on previously identified characteristics. It has become a key component in security governance, credit reporting, and creditworthiness assessments (CWA) that classifies consumers into “high risk” and “high value” groups. In EU law, profiling is considered an automated decision-making process under Article 22 of the General Data Protection Regulation (GDPR), which is in principle allowed but can be objected to by the data subject. According to Recital 24 GDPR, profiling is understood as a decision-making process concerning the profiled individual or for analyzing or predicting preferences, behaviors, and attitudes.

This definition in EU law arguably falls short of the actual phenomenon: profiling is more than just a supplementary process that maps preexisting identities that exist independently of the profiling practice but one that thrives on continuous reconfiguration of technically-mediated identities. An outright “financial identity”—a hidden record of trustworthiness—is constructed by credit bureaus and the individuals themselves. But the aim of the creditworthiness system is no longer the prevention of default, but the expansion of the credit business by repacking and commodifying information and selling customer lists and information. Moreover, competitive pressures force data-processing companies to not merely predict but modify human

---

40. Id. at 18–19.
behavior; it is here where companies can achieve a competitive advantage.\textsuperscript{41}

In the language of systems theory, profiling practices belong to different subsystems, where it fulfills distinct communicative functions depending on the medium involved. What is being communicated are characteristics of profiled individuals. While the aim is always the sorting of individuals into categories, the objectives and patterns of these categories can be manifold. For example, in the economy, based on the medium of money and its reproduction based on the payment/nonpayment code, the goal is to reduce uncertainty about payment capacities. In the case of financial profiling, this is to enable somehow “better” or more efficient economic decision-making.

Algorithmic profiling is a case of algorithmic regulation. Automatic profiling based on algorithmic processes helps to diminish uncertainty by relieving individuals and organizations from the burden of constant meaning-constructing exercises, because the algorithmic logic spares human actors the responsibility of interpretation and evaluation.\textsuperscript{42} These shortened decision-making processes make profiling arguably more efficient because more time-intensive human interpretation and evaluation is being bypassed or made obsolete.\textsuperscript{43} This has to do with a new type of knowledge creation in which data, information, and knowledge are assumed to be the same.\textsuperscript{44} To be sure, the conjunction of algorithms and data, intentionally or incidentally, leads to forms of social ordering through which a prespecified regulatory purpose is achieved, typically to shape, constrain, and coordinate behavior and to manage particular kinds of risk.\textsuperscript{45} In contrast to traditional forms of governance, which are concerned with bureaucratic imperatives and rules, the novelty of this type of algorithmic social ordering lies in its concern with a desirable outcome\textsuperscript{46} that is implemented by a binary mathematical code.

Through its concern for a desirable outcome, profiling practices also

\textsuperscript{41} Shoshana Zuboff, \textit{Big Other: Surveillance Capitalism and the Prospects of an Information Civilization}, 30 J. INFO. TECH. 75, 75 (2015); Kapczynski, supra note 38, at 1489.
\textsuperscript{42} Rouvroy, supra note 2, at 143.
\textsuperscript{43} Antoinette Rouvroy, \textit{The End(s) of Critique}, in PRIVACY, DUE PROCESS AND THE COMPUTATIONAL TURN 143–167 (Mireille Hildebrandt & Katja De Vries eds., 2013) at 147.
\textsuperscript{44} Id. at 147–48.
\textsuperscript{45} Ulbricht & Yeung, supra note 2, at 2.
\textsuperscript{47} A. Aneesh, \textit{Global Labor: Algocratic Modes of Organization}, 27 SOCIO. THEORY 347, 350 (2009) (referring to this type to social ordering as algocracy).
shape social norms and, therefore, have normative effects. Citizens are viewed as entrepreneurs who are empowered to take care of their own affairs thanks to ubiquitous data feedback loops, while the role of the “algorithmic state” is to ensure that reputation becomes a feedback-friendly social net that rewards responsible citizens and protects them from the vicissitudes of deregulation. This implies that social interactions need to be continuously recorded and assessed, erasing whatever difference exists between social domains. As a result, any concern for social justice becomes invisible, because it is “harder to track than the daily routines of the individuals it affects.”

Profiling as algorithmic regulation also institutionalizes moral judgments. They implement a moralized system of opportunities and fair rewards that feedback to people in the forms of regard and disregard. Profiling systems thus implement visions of morality in terms of a person’s integrity, character, and social conceptions of the goodness or badness of one’s numbers. For example, CWAs aim at turning consumers into morally responsible, obedient, predictable, and profitable persons and demand that consumers embrace a regime of continuous inspection and accounting. Based on economic self-awareness, individuals are to manage and actively construct their “financial identities.” Often enough, this goal will be reached through specific and strategic behavioral change.

Last but not least, profiling also exemplifies the expansive tendencies of machine systems. Commercial profiling is based on large-scale data collection and, thus, thrives on and further enables datafication. Datafication as a pathology is the excessive growth of the means of machine communication. The increasing reliance on data in many different systems, such as the economy, is a sign of colonization through the means of machine communication (data). SC is then interested in understanding how datafication can be limited.

50. Id.
52. Marion Fourcade & Kieran Healy, Seeing like a Market, 15 SOCIO-ECON. REV. 9, 10 (2017).
53. Lauer, supra note 41, at 4.
54. Id. at 16–17.
III. The Constitution of the Digital Sphere in the EU

We will now put together the legal framework for profiling EU law within the theoretical framework of SC. According to SC, constitutionalization demands that the legal order of a subsystem performs constitutive functions, that is, formalizes (autonomizes) the communicative means of the subsystem with a view to secure the “self-foundation” of the system (III.A). In a second dimension, SC demands spheres that ensure the possibilities for expressing dissent as a pivotal counteraction to datafication and machine colonization. These spheres can be recognized in risk-based regulation that is based on balancing interests (III.B) and dissent can be expressed both in institutionalized (III.C.1) and the ‘spontaneous’ spheres (III.C.2).

A. Securing the Autonomy of Data

Securing the autonomy of data in EU law language implies its free flow, unimpeded by political or other influences. Recitals 6 and 9 GDPR proclaim the objective of the GDPR to “facilitate data flows” while observing a high level of protection of personal data. The same holds true for the proposed Digital Service Act (DSA) or the Artificial Intelligence Act (AI Act). Recital 52 of the proposed DSA reiterates the dangers of discrimination based on profiling that is inherent in online advertisement and presents a response in terms of providing transparent information and consent for automated decision-making. The focus on transparency underlines the DSA’s concern, included in its Recital 4, with the establishment of “effective and proportionate mandatory rules” providing “the conditions for innovative digital services to emerge and to scale up in the internal market.” The AI Act, based on Article 114 of the Treaty on the Functioning of the European Union, is also directed at ensuring the functioning of the internal market. The proposal is embedded into the core of the digital single-market strategy, aiming at harmonizing rules on safety in compliance with fundamental rights obligations. It aims at ensuring the existence of a European AI market where both benefits and risks of AI systems

55. TEUBNER, supra note 6, at 75.
58. Id. at 1–2, 5.
are addressed at the EU level. To this end, Recital 1 of the proposed AI Act states that the high level of protection of health, safety, and fundamental rights are to be an integral part of the free, cross-border movement of AI-based goods and services.

The establishment of the internal data market is of the utmost concern. Article 20 GDPR, for example, gives data subjects the right to transmit their data to other data controllers. The GDPR’s regulatory aim of establishing a functioning internal market in personal data is at the forefront here. The market should be able to operate with the utmost freedom possible: Recital 13 GDPR even expressively subordinates the protection of individual data to the free flow of data by claiming that the free movement of data requires the absence of restrictions or prohibitions connected with the protection of natural persons and their personal data. In sum, with the aim of the legal framework to safeguard the free flow of data, datafication should be very much served.

B. Limiting Datafication and the Expansive Tendencies of Machine Systems Through Reflexive Law

There are different possibilities to achieve the self-limitation of machines systems through law. Inasmuch as law leaves enough leeway for experimentation and reflexive processes within the (sub-)systems, it is very much in line with SC.

1. Risk-based regulation

The EU legal framework for the data economy is not rife with prohibitions, but the proposed AI Act\textsuperscript{59} foresees a risk-based approach to technologies employing AI and does include prohibitions. Its Title II, Article 5(1)(a) prohibits those AI systems that deploy subliminal, subconscious techniques in order to materially distort behavior if this causes or is likely to cause physical or psychological harm.\textsuperscript{60} Likewise, Article 5(1)(b) proposed AI Act prohibits—in accordance with the Unfair Commercial Practices Directive (UCPD)\textsuperscript{61}—AI systems that exploit

\textsuperscript{59} Id.
\textsuperscript{60} Id. at 43.
vulnerable persons in order to materially distort their behavior. Importantly for our purposes, it also prohibits AI-based social scoring for the evaluation by public authorities, or for their benefit, of the trustworthiness of natural persons based on their social behavior or personality characteristics. In some cases, this scoring leads to detrimental treatment in social contexts that are unrelated to the context of original data collection or where the detrimental treatment is disproportionate to their behavior; this is acknowledged in Article 5(1)(c)(i)–(ii) of the proposed AI Act.62

However, the prohibition of social scoring under Article 5 the proposed AI Act does not include scoring for the purposes of the CWA. This type of profiling is included in the framework for high-risk AI systems in Article 6 Proposed AI Act, which is not prohibited but subject to regulatory oversights and risk management according to Articles 8 et subs proposed AI Act.63 In marked contrast to prohibited AI systems, Preamble 37 and Article 6 exempt the CWA from its scope of application when put into service by small-scale providers for their own use.64 This would imply that small financial providers would not have to comply with the obligations arising under the AI Act for the use of AI in credit scores.

The risk-based approach supports the view that the objective of the proposed AI Act is not to limit the expansion of data or the use of AI. If its concern was the limitation of datafication itself, the proposal would not exempt from its scope the use of AI by small companies but impose the same restriction, including prohibitions, on all commercial users of AI that use big-data profiling.

This lack of prohibition of datafication sits well within SC because of Teubner’s skepticism toward “hard,” state-based regulation. Even though law does play a role in establishing autonomy, SC prefers internal and experimental self-reflection over the prescription of “external” goals exclusively based on political rationality. A risk-based approach allows for experiments and systemic learning, where risks are to be found and assessed in a dynamic way rather than through ex ante regulation.

2. Data Minimization Principle as a Balancing Tool

The norm that could potentially limit the expansive tendencies of data is the data minimization principle (DMP) contained in Article

62. AI Act, supra note 59, at 43.
63. Id. at 45–46.
64. Id. at 27, 46.
5(1)(c) GDPR, which states that the collection and use of personal data should be “adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.” It is, *prima facie*, opposed to any big-data analyses and datafication. The DMP demands that the responsible data processor must justify their need to collect personal data or at least consider using aggregated, anonymized, or pseudonymized data.  

Nevertheless, it is questionable whether the principle is actually designed to limit the expansion of data. Although Article 5(2) GDPR establishes the accountability for observing the data minimization principle on behalf of the data processor, it is unclear whether and how the data processor can actually prove adherence to the principle. The standard for the justification and its relation to the necessity for business purposes remains unclear. In the case of commercial profiling, we must also consider that business models based on behavioral advertising run counter to the DMP by their very nature. If advertising based on behavior and profiling insights are a company’s primary sources of revenue, the company can argue that data processing is necessary for its operational purposes. And if digital companies can justify big-data collection and use by mere reference to some internally defined business purpose, datafication will not be countered. Thus, there is a structural impediment in the design of GDPR rules for the workings of the DMP. It is not interpreted as a prohibition to collect certain types or amounts of data but rather as a kind of balancing tool between privacy interests and business interests (“necessity”).  

Again, this outcome is very much aligned with SC. Being concerned with the control of self-regulatory processes, SC would be incompatible with the idea that data minimization should always prevail. Instead, using the DMP as a balancing tool enables the recognition of different interests through which flexible, experimental reflexive processes can unfold and ultimately lead to system-internal change.

**C. Spheres that Guarantee Dissent: External Pressures**

SC also demands spheres that guarantee dissent. The legal institutions of property and liberty that perform the task of guaranteeing dissent in the political system must be complemented in SC by “reflection centers” within the differentiated society. To this end, SC requires both decentralized organizations and self-regulating

---

67. TEUBNER, *supra* note 6, at 88–89.
institutions on the one hand, and possibilities for individual dissent on the other. Individual and collective rights together with nonlegal media, such as expert knowledge, political power, and social pressure, as well as monetary incentives and sanctions, play a crucial role as external pressures against a “capillary power” that induce change for learning processes. The assumption is that individuals and other societal actors can build up enough external pressure to limit this power so that a balance between different interests is reached. These reflection centers make use of and construct the cognitive openness of systems as external pressures that are the very basis for contestation and, if strong enough, irritation and, as a consequence, change.

1. **Dissent in the Spontaneous Spheres: Public Outrage, Rights, and Consent**

Dissent in the spontaneous sphere implies the politicization of consumer preferences as well as the institutionalization and judicialization of individual rights. Together, these tools are to instigate processes of irritation and self-reflection.

The politicization of consumer preferences means activism, boycotts, public interest litigation, etc. The power of these nonlegal sanctions is based on a twofold belief: first, that the consumers’ readiness to pay is a strong enough tool in the economic system to provoke change and second, that individual rights should and can play a fundamental role in making the spheres of dissent work. Teubner believes that companies adopt their codes as a result of massive learning pressures from the outside. In this sense, SC is in line with what Kennedy has called the third wave of globalization of law, which is largely carried by societal actors and courts. As a result, the internal constitution changes, because it touches on the most sensitive part of the economy: the willingness of consumers to pay. Protests and soft law are leading impulses for changing the corporate code.

---

68. TEUBNER, supra note 6, at 85–86 (referencing Foucault).
69. Id. at 94.
70. Id. at 91.
71. Id. at 95–96.
73. TEUBNER, supra note 6, at 92.
74. Id. at 95.
Institutionalized individual and fundamental rights also play a crucial role. They often provide the basis for public outrage and judicial action. With their horizontal effects against private parties, SC considers constitutional rights to be adequate to counteract reified wills and preferences. The right to data protection of Article 8 Charter of Fundamental Rights of the EU (CFREU) is an example. It is further specified by the GDPR. Chapter III of the GDPR, Section 3 contains a catalogue of rights that allow individuals to express dissent. For example, Article 16 GDPR gives the right to rectify or complete information; Article 17 GDPR contains the right to be forgotten. To this end, individuals have the right to obtain the rectification, completion, or erasure of their personal data. With regard to profiling and automated individual decision-making, the data subject has the right to object at any time to the processing of personal data for the purposes of marketing, under Article 21(1) and (2) GDPR. Moreover, according to Article 22(1) GDPR, every individual has the right not to be subject to a decision based solely on automated processing that has significant effects on their legal situation.

The arguably most useful tool to express dissent is to deny consent to data processing. Articles 5 I, 6 I GDPR establish that data processing is lawful when based on, inter alia, valid individual consent or its necessity for the performance of a contract. Under Article 7(4) GDPR, it must be assured that consent is given “freely,” which demands real choice and control for the data subject. The individual must be able to refuse or withdraw consent without any detriment (see also Recital 42 GDPR), such as, for example, the downgrading of the performance of the service. Recital 43 GDPR mentions that in cases of a clear imbalance between the data subject and the controller, it is presumed that consent is not given freely.

The “free will” requirement is, however, toothless in the commercial

77. Charter of Fundamental Rights of the EU, O.J. C 364/01.
78. GDPR, supra note 39, at 39, 44–45.
79. GDPR, supra note 39, at 37.
80. European Data Protection Board, supra note 67, at 6–11.
81. The requirements for consent were given further shape in Planet49 and Orange Romania. See Case C-673/17, Bundesverband der Verbraucherzentralen und Verbraucherverbände — Verbraucherzentrale Bundesverband eV v. Planet49 GmbH, 2019 E.C.R. I-801 1, 5; Case C-61/19, Orange România SA v. Autoritatea Națională de Supravăghere a Prelucrării Datelor cu Caracter Personal, 2020 E.C.R. I-901 1, 4.
sector. Recital 43 GDPR only considers an imbalance to exist where the data controller is a public authority (or where consent to different data processing operations cannot be given separately) or an employer. Data processing for the purposes of commercial profiling is, therefore, not per se considered a situation of “imbalance.” Similarly, the Proposed AI Act talks of citizen rights when exposed to public authorities. The same holds true considering the “risk of surveillance,” which is only acknowledged for public authorities and employers, but not for the commercial sector, and for the identification of what is “high risk” in terms of the Proposed AI Act (employment and biometric identification).

The Proposed DMA acknowledges the power of so-called gatekeeper platforms (see Recital 3) where contestation powers of others are reduced. But it leaves aside the pervasiveness of an accumulated power that is shared by digital companies, including also non-gatekeeper platforms. In any case, assuming asymmetries only in the public realm and in employment relations or for platforms that fall under the definition of gatekeeper platforms, the DMA excludes the de facto existing power imbalances present in the commercial sector.

Similarly, consent as a legal basis for lawful data processing is inherently undermined by the other legal bases contained in Article 6(1) GDPR, such as the necessity for the performance of a contract. Together with Article 7(4) GDPR, which allows data processing also without consent, this makes for a powerful counteraction to the possibility to express dissent for individuals. In a similar way as with regard to the DMP, if a company can reasonably argue that datafication is necessary for their business goal, which escapes public or supervisory scrutiny, there is no meaningful way for individuals to express dissent and make data processing unlawful. Instead, companies have considerable leeway to define necessity in line with their business models. In general, there does not seem to be an impediment in the regulatory framework to limit expansive purpose or necessity declarations on behalf of the data collector. In this way, Article 7(4) GDPR can undermine the consent basis for lawful data processing, because consent is simply not needed.

It is this alignment of EU law with the interest of digital business practices that insulates digital market activities as much as possible from prohibitions and other regulatory activity, which mainly serves to

---

82. DSA, supra note 58, at 10.
84. GDPR, supra note 39.
85. Id. at 37.
secure the autonomy of data as the communicative medium of machines. It thus enables datafication.

2. *Dissent in the Professional Sphere: Supervision and DPIA*

Under EU law, supervisory authorities are to exercise dissent in the professional sphere. Articles 57 and 58 GDPR include a catalogue of their tasks and powers: for example, they monitor and enforce the application of the GDPR as well as promote awareness among the population. They also act as data controllers, handle complaints, investigate cases, and encourage soft law initiatives. The powers of those authorities also extend to the issuance of warnings or the temporary limitation, definite limitation, or bans on the processing of data, withdrawal of certifications, imposition of administrative fines, or the ordering of the suspension of data flows to a recipient in a third country or to an international organization.\(^{86}\) Especially the latter are adequate competences with regard to the aim of limiting datafication. Bans on the processing of data or the suspension of data flows prohibit the geographical diffusion of data.

For example, just within the last year, several national Data Protection Agencies (DPAs) have suspended data flows to US companies in the aftermath of Schrems II\(^{87}\): the Danish DPA suspended data transfers to Google Analytics in September 2022,\(^{88}\) the Portuguese DPA suspended census data flows to the US,\(^{89}\) and the Austrian DPA decided that the continuous use of Google Analytics as well as the transfer of data to the United States is illegal.\(^{90}\) Most of the complaints underlying DPA action come from civil society organizations, reflecting a unique network of enforcement in the data economy and the coming together of societal forces as external pressure for internal system change.

Moreover, companies are required to undertake Data Protection Impact Assessment (DPIA) when they engage in data processing with new technologies that are likely to result in high risks to the rights and freedoms of natural persons, Article 35 GDPR. Article 35(3)(a) GDPR explicitly mentions profiling as a case in which DPIA are requested, as

---

86. *See id.* at 69.
90. Datenschutzbehörde [DSB] [Data Protection Authority] Dec. 12, 2021, D155.027 (Austria).
it is a case of systematic and extensive evaluation of personal data on which decisions are based that produce legal effects for the data subject.91 Impact assessments can be considered an integral part of SC. In theory, they provide a reflection center where different stakeholders come together and identify common goals and processes to reach those goals. They allow for the exchange of viewpoints and, in the best case, the integration of diverging interests to mutual benefit. Especially because the GDPR rules do not prescribe a certain methodology, they give space to the companies to decide how to comply with the requirements of undertaking a DPIA. While theoretically very fitting with SC, relying on DPIA as a measure of integration of systems is a highly idealized approach. In practice, impact assessments have become a tool for escaping liability. The DPIA, in particular, has been rightly criticized for being too narrow in scope, focusing only on “high risk,” and neglecting, for example, risks coming from anonymous data processing.92

**D. Interim Conclusions**

The EU legal framework on the data economy is fitting for the application of SC. It contains rules for securing the autonomy of data by building and fostering an internal data market and is light on traditional regulation, such as prohibitions and principles. It is the minimal or experimental regulation that still leaves space for spontaneous change through external irritation by economic actors. Commercial profiling is excluded from the application of risk-based prohibitions, and the DMP and consent provisions can be made dysfunctional by arguing in favor of business necessity. Similarly, there are fundamental and individual rights in place, which are to enable the expression of dissent and judicial dissent. DPAs monitor economic activities. This governance style leaves ample space for balancing and bottom-up change.

However, the spheres that are to operate together with minimal regulation by guaranteeing dissent do not fulfill their function of limiting datafication because the framework is tilted toward business interests at the expense of the interests of the data subjects. While the latter have a catalogue of rights at their disposal, the structural power problems involved in expressing dissent with data collection and use for commercial purposes, such as profiling, undermine the fulfillment of

91. GDPR, *supra* note 39, at 53.
their function to limit datafication. Even the DMP is tilted toward business interests because it is inherently limited to what the business deems necessary. This problem is not solved by supervision through data protection authorities. Even though they have been very active in imposing fines and banning data flows to third countries, their decisions are not geared toward limiting datafication and machine colonization. Limiting the geographical expansion of data, as in the aftermath of Schrems II, is not the same as limiting datafication per se. Datafication is a much more substantive and subversive process than the mere geographical diffusion of data.

IV. Questioning the Assumptions of SC: An Internal Critique

The problem underlying this assessment is that the tools that societal actors have at their disposal cannot fulfill their role because, Teubner’s optimism to the contrary, the actors who employ data (the communication means of machines) for their economic activities are too powerful to be challenged. As we will see in this section, their capillary power is unprecedented and remains unmatched by individual rights or social movements, which are conceptually weak and also undermined by a new kind of governmentality that erodes subjectivity, individual autonomy, and free will. Rage against the machine will simply not develop and exert enough pressure. As a consequence, datafication and machine colonization can proceed undisturbed.

In what follows, I submit that it is useful to combine SC with concepts from other theories, such as Habermas’s colonization theory, as well as with insights of the LPE literature. The LPE literature puts emphasis on the unprecedented power asymmetries between companies and individuals. In this way, they shed light on the difficulties of other societal actors to exert the necessary pressure for change. Habermas’s colonization theory, in turn, helps us to understand that the reason for this disenfranchisement has its root in the pathologies that are caused by undisturbed machine colonization. Incorporating these insights means to move our attention toward the question of whether the assumptions of SC can be taken for granted. This leads to an internal critique of SC.

A. Unprecedented Power Asymmetries

There is an unprecedented mismatch of powers in the digital sphere. On the one hand, there is structural digital power and on the other there is weak opposition. As structural digital power arises from the corporate power, not only over precise relations but also over the entire
framework within which actors move and function, even a combination of social movements, public outrage, and individual and collective rights cannot irritate the functioning of the economic system in the digital sphere. Because of a new type of governmentality and data-behaviorism, the very basis for these counterpowers, free will and autonomous action, is eroded.

1. **Structural Corporate Power in the Digital Sphere**

The LPE literature for the digital sphere is almost unequivocal about the new type of power that is localized not only in the state but also in major companies.\(^93\) Benkler defines power as the “capacity of an entity to alter the behaviors, beliefs, outcomes, or configurations of some other entity,” involving the capacity to reshape terms of engagement.\(^94\) He makes it clear that the promise of the open, democratic internet is currently at risk also through behavioral analyses through big data.\(^95\) Big data collection and processing, he explains, create extremely powerful insights on mass populations to only a handful of entities, which then influence users to form beliefs and preferences, follow behaviors, and increase the probability of—mostly commercial—outcomes.\(^96\) Similarly, Zuboff mentions “instrumentarianism power” as the power to manipulate customers’ behaviors according to business goals.\(^97\) Data sets drive a logic of accumulation which Zuboff calls “surveillance capitalism”: it generates revenues from and market control through the data assets that are acquired by ubiquitous automated operations, which have the aim of predicting and modifying human behavior.\(^98\) In this system of informational capitalism, the human experience becomes the raw material that gets translated into behavioral data, which—to the extent not used for feedback loops is “behavioral surplus”—is turned into *prediction products*, which are, in turn, traded as “behavioral futures markets.”\(^99\) Cohen adds platform power, pointing at the development of an access-for-data arrangement: the use of nonnegotiable terms of use contracts and picture-book

\(^94\) *Id.* at 19.
\(^95\) *Id.* at 23.
\(^96\) *Id.*
\(^98\) Zuboff, *supra* note 43, at 75.
\(^99\) ZUBOFF, *supra* note 101, at 8.
contracts of adhesion\textsuperscript{100} to systematically extract data from their users, asserting their nonnegotiable authority over the conditions of access to goods and services.\textsuperscript{101} Companies exercise structural power through contracts in a political economy of private transnational governance by contract.\textsuperscript{102}

While these views focus on the relationship of companies with the data subjects, corporate power over the data relationship is also embedded in a broader, pervasive digital structure. It is Strange’s seminal definition of structural power that is most apt at grasping this phenomenon. She defines structural power as the power to shape and determine the structure of the global political economy within states, states’ political institutions, and economic enterprises, in which professionals have to operate.\textsuperscript{103} Structural power is not only the power to decide how things are to be done but also the power to shape the frameworks within which states, people, or corporate enterprises relate to each other and among each other. This power is relational because the power of each party in a relationship becomes more or less if one party is also determining the surrounding structure of the relationship.\textsuperscript{104} Even though Strange wrote her book well before the rise of the data economy, her idea of power in the knowledge structure is applicable to algorithmic profiling in the data economy: it comes not from coercive powers but rather consent, conferring authority (supposedly) voluntarily on the basis of shared belief systems and acknowledgement that the particular knowledge is important for both the individual and society.\textsuperscript{105}

This structural power is very visible in profiling: the mutually enforcing effects of multiple technologies are designed to influence individual behavior at different stages of the transaction process.\textsuperscript{106} This means that it might be often unclear where to precisely pinpoint the concrete action of manipulation and influence that in the end leads consumers to alter their behavior. The power to shape the behavioral framework for consumers thus lies with the companies that design or

\textsuperscript{100} See generally Friedrich Kessler, \textit{Contracts of Adhesion—Some Thoughts about Freedom of Contract}, 43 \textsc{Colum. L. Rev.} 629 (1943) (exploring freedom of contract and how it should be applied for different kinds of contracts).

\textsuperscript{101} JULIE E. COHEN, \textit{Between Truth and Power: The Legal Constructions of Informational Capitalism} 44 (2019).


\textsuperscript{104} Id. at 25.

\textsuperscript{105} Id. at 118.

\textsuperscript{106} Eliza Mik, \textit{The Erosion of Autonomy in Online Consumer Transactions}, 8 \textsc{L. Innovation & Tech.} 1, 6 (2016).
employ this digital architecture. This power goes beyond the determination of the terms of commercial exchanges and extends to the whole architecture of the digital sphere. The unprecedented power of digital companies is, indeed, a kind of capillary power that “permeates the social body through to its very microstructures.”

2. Weakness of Opposition

The ability of individuals and other societal actors to confront the structural and systemic power of commercial profilers is highly questionable.

First of all, data subjects are either uninformed, apathetic about their privacy, lack the time and resources to actually backtrack the use of their data, or suffer from information overload. One should not forget that, on any given day, individuals are not just confronted with one digital service provider but probably with dozens of data processors that all have their own data terms and conditions. As a result, and notwithstanding the legal rights available to them, individuals are in a considerably inferior position vis-à-vis digital profiling companies and become unable to exercise any meaningful opposition.

Second, individual rights only give a small amount of power that cannot successfully act as a bulwark against privacy intrusions not only because of behavioral issues but also, more importantly, because privacy problems are systemic and relational. Individual rights do not capture the inherently horizontal aspect in privacy protection, which derives from the population-based relations that are construed through data. Adequate protection cannot be achieved through horizontal effect “against social constellations of power,” as Teubner advocates, because horizontal effect does not concern the relation between different data subjects but the relation between data subject and data processor. Besides eroding the capacity for self-formation, datafication also enacts and amplifies social inequality. Strengthening individual rights is simply beside the point if one wants to adequately capture the systemic violations of privacy in population-based relations.

107. Teubner, supra note 6, at 86.
110. Teubner, supra note 78, at 385.
112. For example, Celeste argues in this volume that an Internet Bill of Rights can play a stimulating role to foster debate on how to adapt existing norms and highlighting areas of significant discrepancy between law and social reality.
Third, it is too optimistic to rely on social movements and public outrage to irritate a system sufficiently in order to achieve systemic change. While the readiness of consumers to pay is indeed an instrument to effect change in a business’s orientation or product pallet, until now consumer readiness to pay has not been a driver of real change. For example, on the global level, the offer and purchase of sustainably produced products is still a fringe market, despite continuous growth in the last decades. The problem with the willingness to pay argument in the digital sphere is that it does not necessarily reflect actual preferences. If companies control the access to consumer goods and services, including necessities, any preference is annihilated. Willingness to pay rather facilitates the transformation of privacy into a commodity, which can engender and worsen unequal access to privacy as well as discrimination.113

Fourth, the experimental way of governance based on judicial and supervisory decisions as one way to exert power over companies is maybe too slow. It cannot be easily assumed that extremely big companies, such as Google, will obey rulings and comply with fines based on data protection rules. There have been only a few seminal cases decided by the courts, and it is unclear to what extent companies (primarily Facebook and Google, but also other companies) actually comply with them or whether they continue their business as usual for as long as there are no official complaints about their practices. Chances are high that the latter is the case given the slow speed of court proceedings, the lack of knowledge, the understanding and time of individuals, and the lack of funding and personnel for data protection authorities. After all, the decisions of the supervisory authorities, including fines, were adopted because Google did not comply with the Court of Justice of the European Union Schrems judgments.

This underlines the need to debate whether the combination of the power of individual data subjects, supervisory authorities, and collective interest entities can really lead to the much-needed pressure for change. For example, Teubner understands individual rights as power acts that are collectively aggregated during the formation of a new type of power.114 But even if combined with other loci of power, this counterpower might still not be able to catch up and build up the much-needed pressure against machine colonization through datafication. Machines might just always be one step ahead because they are faster and because their power is structural. In contrast, individual and collective power is dispersed and, similarly to lobbying power,

114. Teubner, supra note 77, at 200.
instrumental (that is related to a specific goal). It can only act within the framework established by the one having structural power. The collective aggregation of rights is not qualitatively the same as structural power. This is where the pervasiveness of structural power comes from.

B. Governmentality and the Pathologies of Datafication

What is more, a new type of governmentality that derives from datafication can make it difficult to identify the “free will” that is to underpin consent and any other criticism or opposition. In this section, I will attempt to give shape to this problem by seeing datafication as a type of colonization of the Lebenswelt that erodes the conditions for the formation of free will.

1. The Pathologies of Datafication

Undisturbed datafication leads to colonization by machine systems. Habermas elaborates descriptions of the problems that come from colonization by explaining pathological deformations in each of the structural parts of the Lebenswelt. Even though, in SC, there is no space for the concept of Lebenswelt, SC does incorporate the Habermasian idea of colonization as a capture by other systems. This idea provides a useful tool to understand the pathologies ensuing datafication. For Habermas, these pathologies are manifestations of crises in the case of disturbed reproduction. There are three: the loss of meaning, anomie, and psychopathologies.

The loss of meaning is a pathology that relates to the cultural dimension of the Lebenswelt, meaning the disturbed continuity between new situations and existing values, which leads to a lack of sufficient interpretative schemes to secure semantic continuity and coherence in everyday life. As a consequence, actors are unable to interpret new situations on the basis of accepted interpretative schemes and the valid cultural stock of knowledge. In algorithmic profiling, for example, honesty and responsibility about the dutiful repayment of credit could be turned into instrumental adherence with the goal of obtaining credit. The entrepreneurial character of “financial identity” and its active

117. HABERMAS, supra note 29 at 215–16.
118. Id. at 212–13.
management exemplify that a “good” financial identity can become an end in itself, detached from the existing moral values of honesty and responsibility.

Anomie, in turn, relates to the main pathological effects inside the socialization dimension of the Lebenswelt. It describes the inability of actors to coordinate in new situations from their inventory of orders because the legitimately regulated social memberships are no longer sufficient and social solidarity becomes scarce. This does not mean that social norms and regulative group identities disappear but that the adherence to social norms is no longer based on solidarity and the interpersonal convictions that those social norms are legitimate. In profiling, this means that the adherence to the social norm of being a “good” debtor is not embedded anymore in a relation between the lender and borrower, which can lead to a certain solidaristic bond between both parties but rather motivated by the managerial attitude adopted toward one’s financial identity.

Lastly, Habermas’s “psychopathologies” describe the pathological deformations in the communicative competences that are necessary to maintain a personal identity. Teubner calls this a mass “violation[] of living [data] subjects’ psychological-physiological integrity” through the uncontrolled accumulation of data in profiles. Datafication in profiling affects psychological-physiological integrity as actors become incapable of maintaining the intersubjectivity of commonly defined situations of action. In order to preserve the personality system, actors need to take recourse to defensive strategies which are, in fact, an impediment to participating in reality-based social interaction. As a consequence, “the strength of the ego” diminishes. This deformation leads to disturbances not only in social interaction but also, importantly for profiling, in a positive self-relation concerning a continuous and coherent life history. Both social reality and inner motivations cannot be assessed realistically anymore, generating fragmented forms of consciousness that hinder an effective use of communicative

120. Id.
123. HABERMAS, supra note 29, at 213.
124. Silva, supra note 123, at 37.
potentialities.125

In systems colonized by data, a digital identity distinct from the actual human identity of individuals emerges.126 In profiling, this is the profiled identity. In creditworthiness, it is the “financial identity.” These digital identities are composed of and constructed by a myriad of proxies, in which individual identity arguably dissolves. The wordplay of Deleuze, according to whom individuals are disassembled into “dividuals”127 underscores this: individuals are no longer the smallest unit in society; they are broken down into small data points, deconstructed into different roles which they play in more or less random communities. However, this disassembled self is a “probabilistic subject” that is not the same as the actual, experiential subject.128 Being part of one community of “dividuals” who share a certain criterion in one field does not mean that all individuals assembled under this community have anything else in common. Outside of the digital realm, but in real, experienced life, individuals are more than the sum of “dividuals.” The fractured, digital pictures that are assembled through data proxies sit uneasy with the individual need to feel complete and might even prevent or restrict individuals from relating their identity to them.129 As Habermas argues, the basis for coherent life stories and truthful inner motivations is eroded.

2. The Erosion of Free Will Through Governmentality/“Data-Behaviorism”

As a consequence of these psychopathologies, we must question the “free will” underlying any consent in the data economy. This goes beyond merely behavioral issues, such as the ones identified by Acquisti et al. They show that it is virtually impossible for individuals to provide meaningful, voluntary consent to the sharing and processing activities, which algorithmic regulation entails.130 The behavioral issues in this

125. Id.
128. Rouvroy, supra note 2, at 8.
129. See De Vries, supra note 40.
regard are well known, mainly the privacy paradox and information overload that apply to data policies as standard terms. The privacy paradox is fueled by a new underlying logic of exchanging privacy for convenience and perceived efficiency. While these behavioral issues could maybe be addressed—for example, with shorter data policies—the lack of free will goes deeper: the mere will for possible dissent is undermined because of the inbuilt governmentality effects of profiling algorithms. As stated above, algorithmic governmentality, by its very design, avoids all types of confrontations. It is the shortening of human decision-making processes that increases economic efficiency that is the very reason for engaging in profiling.

Rouvroy has termed “data-behaviorism” the new regime of truth that erodes uncertainty as a precondition for the exercise of critique. A system of anticipatory conformity is established in which mental agency and self-possession of anticipation is submerged into a new kind of automaticity. This means that the preemption of the space for critique, dissent in Teubner’s SC, is also inbuilt into the very algorithmic architecture by design. The preemption of critique is the direct consequence, if not even part of the raison d’être of profiling: any possibility for critique would make the decision-making process more complex and lengthier and, thus, less efficient. As Zuboff says: “[i]n this world of no escape, the chilling effects of anticipatory conformity give way as the mental agency and self-possession of anticipation is gradually submerged into a new kind of automaticity.” Here, “contract and the rule of law are supplanted by the rewards and punishments of a new invisible hand.”

Under conditions of governmentality, consent becomes a mere reflection of discounting mechanisms individuals are subject to when

134. Yeung, supra note 1, at 10.
135. Rouvroy, supra note 2, at 149.
136. Id. at 146.
137. Zuboff, supra note 43, at 82.
138. Id.
139. Id.
wanting or needing access to certain goods and services, as is usually the case with profiling. Subjectivization might not be obvious—it’s indeed “spectral” as Rouvroy states—but it exists. This subversive change of mental agency hollows out any attempt to empower data subjects individually. For this reason, from the point of view of counteracting datafication, proposals, such as empowering individuals by allowing payment with data or creating data property, are also beside the point. If private autonomy is undermined through a lack of privacy and the subjectivization effects of profiling, then any legal construction that is based on the exercise of this private autonomy will not hit its mark. Rage against the machine will simply not happen in this context, simply because there is no space for rage to develop.

V. MAKING SC WORK: TILTING THE POWER

In order to make SC functional in the digital sphere, the existing rules concerning high-risk AI, DMP, and consent should be underpinned by a new concept of structural, digital vulnerability. In this section, I aim to show that this concept could enable stricter interpretation of the rules in order to orient them toward counteracting corporate power and help to abate the extreme power asymmetries.

A. Conceptual Underpinning: Structural Digital Vulnerability

The conceptual underpinning for a change of interpretation of EU law rules on datafication could be achieved by the concept of “digital vulnerability.” This concept was proposed for the EU legal order by Helberger et al. and denotes the lack of power or ability to affect the decisions, desires, and behavior of the consumer in ways that the consumer and data subject cannot prevent. It mirrors the “industry’s relentless search for experimental and creative digital marketing practices that seek to ‘optimize’ consumers’ behavior.” Through these practices, digital companies identify and specifically target

144. Id.
vulnerabilities for their commercial gains. This insight acknowledges inherent and situational sources of vulnerability as well as its "architectural" character. Digital vulnerability is more than just a situation or an "unfortunate by-product" of economic activity in the digital sphere, but rather it is deliberately created by it. It is based on people's ongoing involvement in various digital markets and services that make them increasingly vulnerable to manipulation because the longer the relationship between a consumer and a digital service or app persists, the more the service or app establishes a position of power as a result of increased knowledge about its users.

This conception of vulnerability goes beyond the ideas of vulnerability in the current EU legal framework. Most prominently, Article 5(3) UCPD employs a definition of vulnerability but circumscribes vulnerability to specific (groups of) individuals who display certain characteristics. In contrast, structural digital vulnerability acknowledges the insights of the LPE literature about the extreme power imbalances and its consequences for individual autonomy.

Structural vulnerability could support SC to reach its potential as a metatheory of coregulation. As the other side of unprecedented company power, it could justify imposing more obligations on companies on the one hand and higher barriers for the assumption of autonomy of individuals on the other. More specifically, it underpins a broad and strict interpretation of the DMP and the concept of free will. Coupled with strong enforcement, maybe in expediated court proceedings, it could maybe tilt the power imbalances that currently dominate the digital sphere.

**B. Profiling as Prohibited AI**

Acknowledging the structural vulnerability of data subjects could justify the extension of the risk approach of the AI to commercial profiling, such as CWA, be it the prohibited or regulated AI. Putting emphasis on the structural vulnerability of data subjects in any digital context makes it clear that the distinction between social and

---

145. *Id.*
146. *Id.* at 18.
147. *Id.* at 19.
148. Helberger et al., *supra* note 140, at 20. I do not agree, however, with Helbereger et al's claim that people move in and out of states of vulnerability. Especially with regard to algorithmic profiling, I think that the structuredness of the power relation with the company is so pervasive that vulnerability is omnipresent in the digital sphere.
149. UCPD, *supra* note 61.
commercial profiling, which is included in Article 5(1) AI Act, is untenable. The privacy intrusion is the same. The pathologies are the same. If an outright prohibition in line with Article 5(1)(a) of the proposed AI Act is considered disproportionate, it should at least be considered a regulated AI. Article 5(1)(b) of the proposed AI Act already includes the concept of vulnerability. However, instead of focusing on the discriminatory effects of those AI applications, structural digital vulnerability allows moving toward an interpretation that is concerned with the formation of material conditions for action rather than their effects.

C. Strict and Broad Interpretation of the DMP

The concept of structural vulnerability can help to interpret the DMP in a broader and stricter way. Such interpretation should be the basis for the interpretation of the scope of any rights and obligations.

First of all, interpreting the DMP broadly should imply considering ruling out advertisement-driven business models for which large amounts of data are accumulated and then analyzed for profiling purposes. Data minimization must be achieved through the bench in all aspects of the digital sphere in order to address the pathologies of datafication. Second, the authoritative control of the objectives of the data collection and processing by digital companies should not be off limits. This requires the disclosure of business models for control by an authoritative institution, for example, DPAs. In a digital world, in which almost all businesses capitalize on user data by selling that data to advertisers, it is too easy of a way out for businesses to simply argue that the data collection is needed for this business purpose. While this approach that balances different interests respects the concern for self-regulation, we have seen that the DMP is structurally inhibited because of the possibility to argue for the necessity. The almost mindless or automatic connection of gratuity of the service with the need for an advertised-based business model must be questioned, if the DMP is to be taken seriously.

Of course, the broad interpretation of the DMP runs counter to the current raison d’être of the EU, which still understands itself as—and is probably constitutionally limited to—a market-building entity. However, if datafication and its pathologies are to be taken seriously, as SC demands, datafication itself must be strictly limited, controlled, and even rolled back. The DMP already exists; it just needs to be interpreted so as to be able to serve its very own purpose.
D. Taking Free Will and Consent Seriously

In order to take the freedom of will in the data economy seriously, both the theoretical and legal framework must acknowledge the insights of the LPE literature. First, this requires acknowledging that individual consent is simply not “freely given” in commercial settings. Structural vulnerability permits this interpretation. Again, the current distinction between a social or public sphere, where free will is supposedly absent, and a commercial sphere, where free will is prima facie assumed, seems artificial and is untenable from the viewpoint of a structurally vulnerable individual data subject.

Second, in order to reinforce the importance of free will in the digital sphere, consent should be the main legal basis for data collection. The legal basis of the necessity for business purpose, in turn, should be an exception that should be interpreted narrowly. In this way, business necessity does not become a blank permit, but companies have to fulfill much higher thresholds of justification to overrule consent as a legal basis.

This goes beyond a mere application of the EU consumer law framework, the Unfair Contract Terms Directive 93/13 (UCTD) precisely. The UCTD deals with contracts of adhesion, where there is arguably no free will because those contracts are not negotiated but offered on a take-it-or-leave-it basis. Nevertheless, the UCTD does not aim at the establishment of free will ex ante but rather the prevention of the use of standard terms that are unfair as a possible consequence of the lack of free will. The approach advocated here would not dismiss the UCTD as a useful instrument in the control of unfairness in the digital sphere, but it would go beyond its current text and interpretation because it is concerned with enabling the formation of free will ex ante as the very basis for consent.

VI. Conclusions

The concern of this contribution is the colonization of the data economy by machines. I try to understand to what extent Teubner’s theory can account for the emergence of a SC in the digital sphere with norms that do not only foster a data market but also limit datafication as a pathological deformation of communication and, thus, counteract and roll back colonization.

151. See Kessler, supra note 104.
Taking the example of consumer profiling, I show how economic processes based on big-data analyses are an example of algorithmic governance that incorporates moral and normative standards. As an arguably efficient decision-making tool, it abbreviates decision-making by cancelling the possibility for critique and contestation. It is here where the “digital constitution” should intervene by opening the space for putting limits to datafication and for expressing dissent.

SC provides a useful lens for grasping processes of datafication, as it allows for a problem-oriented view and helps to identify datafication in itself as a problem. Seeing datafication through the lens of SC (and systems theory) shows that the danger of colonization does not only stem from the economy (or law) but also from other types of systems, namely machine systems. As machine systems operate according to binary rationales, which go hand in hand with the rationalization efforts of the economic system with its emphasis on economic efficiency, economic systems and machines have boosting effects on each other.

In general, one can argue that a SC in Teubner’s vein is emerging. EU law fosters an internal data market while being light on traditional regulation of the commercial sphere in terms of prohibitions and principles. At the same time, it establishes a system of enforcement, in which individual rights, supervisory authorities, and judicial action intertwine, thus arguably making space for the critical contestation that is crucial to SC because it can amount to such high pressure that generates internally processed change.

However, the theory of SC runs into difficulties when it comes to the possible solutions to the problems of datafication and counteracting machine colonization. More specifically, the spheres for dissent that are to operate together with minimal regulation do not fulfill their function of limiting datafication because the framework is biased in favor of business interests at the expense of the interests of the data subjects.

The reason for this inoperability of the EU SC in the digital sphere is its underlying assumptions, which are simply not met. Teubner voices a very optimistic view on the capacity of experimental co-governance to—admittedly “with a bit of luck”—bring about an equilibrium between functionally differentiated systems by supporting their autonomy and suppressing their expansive tendencies through societal counterpowers. Respecified constitutional rights are to act as a bulwark against social constellations of power. However, these counterpowers can only emerge as effective dissent when the conditions for the formation of free will are met. But in a digital world, where insights into personal motivations are used for corporate financial gain and where “data-behaviorism” prevails, one can simply not hope for this bit of luck. Pathologies emerge, which Habermas describes. Free will is eroded, and, therefore, also the private
and individual autonomy to express dissent or to exert any type of meaningful pressure for the limitation of power in the digital sphere.

In order to let go of optimism and make SC work in the digital sphere, I submit that the insights of LPE are especially useful. They can help to design a more robust legal framework to tilt back the power and achieve greater balance between corporate and individual interests. The concept of structural digital vulnerability of individuals can act as a counterweight to the underlying structural power of companies. Basically, the more power companies have, the more obligations they should have and the stronger legal protection and empowerment for individuals should be. Structural digital vulnerability can help to interpret the DMP in a stricter way, include commercial profiling into the catalogue of prohibited high-risk AI applications, and strengthen the requirements of “free will” and consent. Taken together, this would allow for a balancing of interests, with the structurally more powerful parties facing more evidentiary burdens.

This analysis is of course limited. Most importantly, I did not look into all the dimensions of SC. Therefore, the question whether a full-fledged constitution has emerged remains subject to further research. What remains clear though, I hope, is that in order to make SC work, legal rules are needed to tilt the power imbalances in the digital sphere toward structurally vulnerable consumers, in order to enable them to exert high pressure against machine colonization and the expansion of datafication: rage against the machine.
Digital Monetary Constitutionalism: The Democratic Potential of Monetary Pluralism and Polycentric Governance

ROXANA VATANPARAST*

ABSTRACT

Focusing more on governance and democratic potential than financial potential, this paper explores how pluralism of digital money and polycentric governance may provide an opportunity to embed values that might otherwise be overlooked or not valued in market societies. Digital money affords diverse forms of democratic experimentation with institutional and technological designs to effect distributive and political transformations for marginalized communities. Examining two case studies, including digital fiat currency that has the privacy-preserving features of cash and promotes financial inclusion and digital money built by and for stateless populations utilizing blockchain technology, the paper argues that pluralism of digital currencies using public and not-for-profit institutional architectures to serve diverse publics’ interests has more democratic potential than forms of digital money driven by profit and extractive motives. To effectuate this democratic potential, however, attention must be paid to the tensions of digital exclusion and governance by experts.

Key Words: Digital Money, Monetary Design, Pluralism, Polycentric Governance, Digital Governance

* Assistant Professor of Law, Capital University Law School; Fellow, Transatlantic Technology Law Forum at Stanford Law School. I would like to extend my gratitude to Angelo Golia, Gustavo Prieto, Daniel Rohde, Thomas Streinz, and Gunther Teubner for their insights on prior drafts. I would also like to thank participants of the Max Planck Institute for Public Law and International Law’s Workshop held in June 2022 on “Digital Constitution: On the Transformative Potential of Societal Constitutionalism” for helpful comments, as well as the Harvard Law School librarians that helped me access sources during an independent research stay in Summer 2022. The usual caveats apply. Contact: roxanav41@gmail.com.
I. Introduction

Conversations around digital money have gained increased traction among scholars, policymakers, and the wider public in recent years with the proliferation of digital currencies built using blockchains. Bitcoin, the first such digital currency built on blockchain technology, was launched following the white paper of a person or persons writing under the pseudonym of Satoshi Nakamoto in 2008, initiating the launch of several other blockchain-based digital currencies or cryptocurrencies, including Ethereum. These and other digital currencies built on a blockchain were distinct from prior forms of digital money not only because of their peer-to-peer qualities but also because of their distributed consensus mechanisms and the possibility for self-executing code and decentralized autonomous organizations (DAOs) that enable social coordination and self-governance for communities. They enabled participants and stakeholders to engage in decision-making and governance of the networks based on, and mediated by, the networks’ underlying code. These forms of digital money and cryptocurrencies have therefore raised conceptual and normative challenges to traditional understandings of money, the global monetary system, monetary sovereignty, and issues of governance, privacy, and distribution.

With the development of blockchain technologies, digital money has taken the form of a multitude of cryptocurrencies and stablecoins, which

1. The term “digital money” or “digital currency” is used in this paper rather than “virtual currency” because virtual currency is often defined in a way that excludes fiat currencies. For example, the United States Internal Revenue Service defines virtual currency as “a digital representation of value, other than a representation of the U.S. dollar or a foreign currency (“real currency”), that functions as a unit of account, a store of value, and a medium of exchange.” See I.R.S. Notice 2014-21, 2014-16 I.R.B. 938, https://www.irs.gov/pub/irs-drop/n-14-21.pdf (last visited Jan. 6, 2023). See also Benjamin Geva, Disintermediating Electronic Payments: Digital Cash and Virtual Currencies, 31(12) J. INT'L BANKING L. & REG. 661-674 (2016) (defining “virtual currency” as “digitally and privately issued.”) Moreover, while digital forms of money and electronic banking and payment methods have existed for decades, this article is using the term digital money to refer to virtual currencies, cryptocurrencies, CBDCs, and digital fiat currencies in the contemporary context. The article will differentiate between the terms where relevant.

2. See Finn Brunton, Digital Cash: The Unknown History of the Anarchists, Utopians, and Technologists Who Created Cryptocurrency (2019) (providing a longer history of some of the researchers and technologists that were developing public-key cryptographic technologies to convert money into a digital format as early as the 1970s.)


some jurisdictions accept as legal tender. Cryptocurrencies are digital currencies issued by private parties on a blockchain and denominated in its own unit of account. These are often issued on a permissionless distributed ledger technology (DLT) (e.g., Bitcoin and Ethereum), though some cryptocurrencies maintain more centralized forms on a permissioned blockchain. Cryptocurrencies may also be referred to as tokens, especially when holding the underlying asset is accompanied by voting rights in governance decisions.

Digital money has also taken more centralized forms, such as private currencies developed by digital platforms (e.g., Facebook’s Diem, formerly Libra) and proposals for public digital money, such as Central Bank Digital Currencies (CBDCs) and digital cash or digital fiat currency (DFC). A CBDC is a central bank deposit in digital form on a blockchain that tends to be in centralized form. CBDCs can be designed in a variety of ways, depending on policy aims. CBDC design choices include access, anonymity, availability, transfer mechanisms, and interest-bearing characteristics.

Digital cash or DFC are similar to e-coins in that they are pegged to fiat currency or are legal tender themselves, but they are backed or managed by governments, and they are not necessarily on a blockchain. Since transactions are not placed on a blockchain, they have privacy-preserving features similar to physical cash, with the added benefit of enhancing financial inclusion for the unbanked and underbanked.

These technological developments have brought about changes in lexicon about not only the technologies but also regarding broader conceptions of money, and the role of banking and monetary institutions. While often used interchangeably with the term “cryptocurrencies,” “virtual currencies,” or “digital currencies” in scholarly literature and policy briefs, digital money as used here refers to both those private and public digital currencies that are built on a

---


7. See generally Ying Huang & Maximilian Mayer, Digital currencies, monetary sovereignty, and U.S.–China power competition, 14 POL’Y & INTERNET 324 (2022) (The project was later canceled. In 2022, the firms developing Diem sold the assets to the financial services firm Silvergate.)

8. See Adrian and Mancini-Griffoli, supra note 6 at 61.

blockchain, as well as digital cash or DFC not built on a blockchain.

Digital money has also raised several issues and conceptual challenges to the global monetary and payments systems, monetary sovereignty, and issues of governance, privacy, and distribution. At the global level, digital money has the potential to challenge the global monetary architecture that has existed since the 1970s, with the US dollar being the primary reserve currency. It not only raises issues concerning privacy and monetary sovereignty but also questions of global distribution and (un)democratic governance as it is regulated by multiple fragmented (and at times conflicting) governance regimes and normative orders, including regulations and technical standards, and mobilizes expert knowledge.

While money is often described as a store of value, medium of exchange, and unit of account, money also has important governance and distributive implications. It is a constitutional project, a mode of ordering and governing. According to Desan, design choices for money have implications for exchange, value, and distribution, not only for the medium of money itself but also for broader society and the economy.

Moreover, scholarship in Science and Technology Studies (STS) has shown that design choices in technologies can have effects on how they are taken up in society and that technologies are political and embed and shape social values. This paper will consider the design choices that can help make digital money more attuned to the needs of democratic publics and society. With regard to money and digital money in particular, design choices can have a powerful impact on governance, institutions, and policy aims.

Drawing from societal constitutionalism (SC), this contribution argues that designing digital money toward democratic ends means having a wide range of monetary designs that enable polycentric and


13. Design choices here refer to both technical and institutional design choices.

14. Polycentricity has been defined as “a social system of many decision centers having limited and autonomous prerogatives and operating under an overarching set of rules.” Paul D. Aligica & Vlad Tarko, Polycentricity: From Polanyi to Ostrom, and Beyond, 25
informal governance and redesigning banking and monetary institutions to enable the pursuit of those ends in line with democratic values, rather than an individualistic paradigm. Digital money affords particular forms of informal governance that can reflect the values of the communities that shape them, as well as possibilities for privacy and financial inclusion for publics. It should be designed in ways to promote these affordances for the public interest.

There are no universal prescriptions in this article, since democracy is a culturally contingent and “essentially contested”\textsuperscript{15} concept that is enacted in a variety of socially and historically situated ways. Nevertheless, for the purposes of this article, the concept of democratic design means one attuned to the needs of communities, and one that is designed for the purpose of providing social welfare and serving public ends, rather than private ends (such as providing speculative value for private investors or opportunities for maximizing individual and institutional profits).

In challenging the monetary system, as well as theorizations and understandings of it, digital money provides an opportunity to imagine and implement new monetary models that are not premised on Lockean individualist profit maximization, but rather to be used for public ends and social welfare, democratically decided. This means thinking about whether digital money can be designed to democratize work, housing, and social ecological welfare, as well as considering what democracy is and ought to mean in the digital context, when it is mediated by data-driven technologies. Designing digital money to provide socially beneficial functions could mean attributing more value to the money when it is used for socially beneficial functions. These socially beneficial functions can be decided upon by a community, particularly considering the needs of the most vulnerable groups. These groups include informal workers, low-income workers, refugee populations, communities most affected by the effects of climate change and environmental degradation,

the unbanked, and other groups that have been historically disadvantaged by banking and financial institutions to great detriment, with effects that persist today.\textsuperscript{16}

In other words, digital money could be a pluralistic, polycentric governed tool for advocating and effectuating social and distributive justice. Thus, rather than the value of digital money being based on exchange rates with other assets and fiat money, we can rethink monetary value as a matter of socially beneficial value. While conventional money is often thought of in dominant strands of economic theory as a tool for allocating scarce assets and labor, digital money grants an opportunity to rethink this and other conventional ontologies of money altogether. This also creates a new meaning for monetary sovereignty outside the confines of the international financial and economic system institutions. It may offer an opportunity, however slight, for experimentalism that might not be available otherwise with conventional money, including increased investment in education, health, social welfare, infrastructure, and other services. It offers new opportunities for not only digital public cash and financial services for the unbanked but also for using money as a tool for collective action.\textsuperscript{17} Even if not actualized, the potentiality of these technologies to help rethink the status quo of governance and distribution is valuable in itself.\textsuperscript{18}

This paper will explore two examples to illustrate the benefits of embracing monetary and normative pluralism through polycentric governance and design of digital money in different contexts: (1) e-cash or DFC in the context of the bill proposed in the United States as The ECASH Act and (2) R-Coin, a cryptocurrency being developed by and for the stateless Rohingya population as well as other blockchain technology projects being developed for political dissidents.

To effectuate its democratic potential, digital money must move away from profit-oriented forms of money to counter the tendency of private money creation to seek ever greater extraction, of both labor and natural resources, for the sake of increasing profits.\textsuperscript{19} The examples

\textsuperscript{16} See Mehrsa Baradaran, How the Other Half Banks: Exclusion, Exploitation, and the Threat to Democracy 1, 139 (2015).
\textsuperscript{17} See Giacomo Bazzani, Money as a Tool for Collective Action, 13 Partecipazione e Conflitto 438 (2020).
explored in this paper are tellingly not those that are promoting profit-seeking interests, which can ultimately lead to fraud and financial manipulation, as seen in the recent collapse of FTX, among other cases. The examples explored here show that the technologies can embed different sets of values in their architecture when developed by public and not-for-profit actors, indicating just some of the ways that institutional architecture matters as much as technological architecture in promoting values oriented toward communities and for the public good.

No democratic futures can be brought about through technological innovation alone. Rather, technology’s embedded politics and broader interaction with social and legal institutions can collectively promote nonmarket-based value systems oriented to diverse publics’ interests, even when those publics and communities are not recognized by any state. Moreover, digital money’s democratic potential may be limited by digital exclusion and governance by experts. Attention must be paid to whether digital money might provide an opportunity to deconstruct the binaries of democratic politics versus rule by expertise and how to address digital exclusion in this context.

II. DEMOCRATIC QUESTIONS RELATING TO DIGITAL TECHNOLOGIES AND DIGITAL MONEY

A. What is Meant by Democratic?

SC is a means of analysis and critique, but also a normative framework for addressing some of the problems associated with sociotechnical systems and the liberal constitutionalist and proceduralist orientation that lies at the core of many current scholarly and policy projects in the digital context. As a mode of critique, SC accounts for the plurality of normative orders, including those which derive from nonstate actors and systems that develop “constitutional qualities, notably in the transnational economy and the digital sphere.” It accounts for the distinct normativity of functionally

22. Id.
differentiated systems, such as economy, science, and religion, as well as the digital sphere, and “strives to democratise autonomous spheres of a functionally differentiated society, beyond the institutions of state politics.” Once one begins to ask questions around what is distinctive about digital normative orders, one might better understand its interconnections with other social systems and what kinds of institutional frameworks or responses might be available to address them.

According to SC, increased functional fragmentation has shaped transnational “law without a state” or global private legal regimes, such as *lex mercatoria* and in the digital context *lex informatica* and *lex cryptographia*, or the “rules administered through self-executing smart contracts and decentralized (autonomous) organizations.” This law without a state, however, complicates how responsibility is attributed, when fragmented and diffuse social systems and normative orders are not adequately accounted for in traditional conceptions of fundamental rights or human rights invoked against states or state actors. This is particularly true in the digital context, dominated by normative orders produced by private actors and the “code is law” or “law is code” system.

To address some of the democratic limitations of transnational regimes theorized in societal constitutionalism, Teubner articulated a conception of democracy where self-contestation is institutionalized. Self-contestation in this context refers to responsiveness to external “irritants” and institutionalized sites of internal dissent. In this conceptualization, self-contestation must be plural to reflect the “epistemic diversity among issue-specific transnational regimes.”

---

23. *Id.*
24. *Id.* at 5.
32. *Id.* at S14.
33. *Id.* at S14.
In this view then, if different forms of digital money are themselves oriented to serving a variety of specific issues, then they might each form separate transnational regimes with their own modes of self-contestation. However, these regimes are not entirely self-contained—there is also interaction with external factors and institutions that also shape them and their democratic legitimacy.

While some have utopian, optimistic views of the democratic potential of digital technologies, this vision is subject to the critiques of increased datafication, loss of privacy, surveillance, cybersecurity risks, lack of access to technology, information filtering and echo chambers, and potential for political manipulation and undermining democracy vis-à-vis digital technologies and platforms. For all the promises of the democratizing potential of new technologies, these promises often turn into antidemocratic failures, as has also been noted regarding blockchain and cryptocurrencies, such as Bitcoin. However, the democratic potential of digital technologies depends on how they are designed and how they are taken up and reshaped by communities, laws, institutions, and social order.

One of the crucial pieces in the context of digital money, moreover, is how the members of a system are defined and socially, juridically, or constitutionally related, as well as how the association of a polity is defined. The configuration of simultaneously siloed and overlapping communities further complicates this vision of democracy as a normative ideal. This is especially the case when dealing with multiple levels of social and technological decision-making processes—whether at the level of formal voting systems, participation and representation in society or vis-à-vis technology, or in the context of digital monetary constitutionalism.

One of the central questions therefore is: who is the demos of the various forms of digital money? The binary between the governors and the governed is not as clear in this context, which provides further opportunities for new ways of imagining democratization of digital money. This is something that Teubner points to, arguing that a plural notion of demoi and several public spheres would make more sense in this context, rather than more traditional Habermasian notions of the

34. See Bryan Ford, Technologizing Democracy or Democratizing Technology? A Layered- Architecture Perspective on Potentials and Challenges in DIGITAL TECHNOLOGY AND DEMOCRATIC THEORY 274, 275 (Lucy Bernholz, Hélène Landemore, & Rob Reich eds., 2021) (examining the promises, failures, and challenges for creating a system of digital democracy).

35. See Teubner, supra note 31 at S15–S16 (On similar questions with regard to transnational self-contained regimes more generally).

36. Id. at S16.
public sphere and democratic deliberation\textsuperscript{37} that do not fit squarely with transnational regimes, such as those afforded by digital money. These \textit{demoi} could be distinguished as “regime ‘participants’: members, rule addressees, [and] affected outsiders.”\textsuperscript{38}

If we consider money as infrastructure,\textsuperscript{39} we might also consider these \textit{demoi} as “infrastructural publics”—“people brought together by being affected by the same infrastructures”—in addition to legal publics and normative publics.\textsuperscript{40} These overlapping and plural publics may better accommodate political contestation than a singular notion of the public.\textsuperscript{41} These infrastructural publics are often thought to be divergent from legal publics.\textsuperscript{42} The gap between who gets to govern, make, shape, and resist infrastructures ought to be accounted for and integrated if there is to be a democratic theory that adequately captures these relationalities.\textsuperscript{43}

The potential for both internal and external contestations and resistance at multiple levels of \textit{demos} and multiple regimes renders these regimes sites of politics and politicization in ways that are not available with formal political institutions. As each regime has its own communicative and knowledge patterns, “counter-expertise becomes the relevant counter-power” in ways attuned to the specific patterns of each regime.\textsuperscript{44} The relationship between these regimes can also be a site of democratic contestation to avoid the hegemony of any one regime over others, especially when a regime’s constitution reflexively contemplates procedures for “regime collisions” or conflicts with other regimes.\textsuperscript{45}

Moreover, money holds potential as a site of resistance against surplus production tendencies where “new modes of economic integration (redistribution, reciprocity and householding)” can be created toward “non-profit-oriented modes of economic organization.”\textsuperscript{46} Teubner proposes that “non-surplus-oriented institutions” need to be oriented toward “social protection” of “people, society and nature.”\textsuperscript{47}

\begin{thebibliography}{99}
\bibitem{37} See \textit{generally}, \textsc{Jürgen Habermas, Between Facts and Norms} (1996) (referring to Habermas’s argument for informal deliberation in the public sphere and representation in a democracy through representatives in formal decision-making processes).
\bibitem{38} See Teubner, \textit{supra} note 31 at S16–17.
\bibitem{40} Benedict Kingsbury & Nahuel Maisley, \textit{Infrastructural & Legal Publics and Publicness}, 17 \textsc{ANNU. REV. LAW SOC. SCI.} 353, 358 (2021).
\bibitem{41} Id.
\bibitem{42} Id. at 354.
\bibitem{43} Id. at 361, 367.
\bibitem{44} Teubner, \textit{supra} note 31 at S26–S27.
\bibitem{45} Id. at S22–S25, S29.
\bibitem{47} Id. at 516.
\end{thebibliography}
These institutions need not only derive from the state or public institutions. As he states, “collective actors should be created which are experimenting with new democratic forms outside institutionalized politics.”

Designing digital money to decentralize decision-making through the plurality of its forms, polycentric governance, and policy-specific regimes might afford more democratic promise than more conventional forms and conceptualizations of money as fiat currency backed by the state. By multiplying the sites of transnational decision-making and contestation, pluralism and experimentation in a dynamic process might offer more democratic promise than more centralized and stabilized forms of global governance.

The next sections will contextualize the debates around democratizing components of today's digital economy, such as personal data, algorithms, and artificial intelligence, as well as blockchain governance, and discuss some recent debates around democratizing money in contemporary economies. It will then analyze global dimensions of digital money.

1. Democratic Data, Algorithms, and Artificial Intelligence (AI)

In recent years, there has been a growing literature and public interest in models for conceptualizing data democratically, democratizing digital technologies, and promoting democratic values in technological innovation.

In the context of personal data, Viljoen has argued for a relational approach for governing personal data because of the collective interests that data analytics touch upon. In Viljoen’s approach to democratic data, she highlights the importance of the vertical and horizontal relations that data make that permit population level insights and analytics that constitute the data economy. In contrast to current data governance and legal models that focus on individual proprietarian or dignitarian rights in data, she elevates these relations as the subject of concern, discusses institutional proposals that align with this relational approach for governing data, and promotes avenues for promoting

48. Id.
49. Id.
52. Id.
public use of data for beneficial purposes.\textsuperscript{53}

Okidegbi has argued that the exclusion of marginalized communities in the governance of pretrial algorithms reproduces social, economic, and racial inequalities.\textsuperscript{54} These marginalized communities are excluded from democratic participation at multiple levels in society—not only in the governance of pretrial algorithms but also in broader democratic participation in society.\textsuperscript{55} Input from these communities alone would not solve the democratic deficit and inequalities associated with the use of these algorithms, with the law and policy that is shaped by them, or with the way that criminal law facilitates exclusion of these communities from broader democratic participation by “diluting their political, economic, and social power . . . .”\textsuperscript{56} She argues that a participatory model is not enough due to power differentials\textsuperscript{57} and that power shifting is more apt to produce changes that address power dynamics in policy and decision-making that tend to be technocratic and antidemocratic.\textsuperscript{58} She proposes two models for this power shifting: inclusion of marginalized communities in decision-making power as well as representation of these communities in policy making institutions.\textsuperscript{59}

Regarding artificial intelligence, scholars have argued that participatory AI holds some promise to acknowledge, include, and empower marginalized communities in developing technologies that can promote public interests.\textsuperscript{60} However, these are not substitutes for the role of public institutions in a democracy.\textsuperscript{61}

Using insights from these analyses, this paper will consider ways of conceptualizing digital money democratically, as well as ways to engage marginalized communities in decision-making processes and include them in institutions from which they have been historically excluded.

\textsuperscript{53} Id.
\textsuperscript{54} Ngozi Okidegbe, \textit{The Democratizing Potential of Algorithms?}, 53 CONN. L. REV. 739, 739 (2022).
\textsuperscript{55} Id.
\textsuperscript{56} Id. at 765–66.
\textsuperscript{57} Id. at 773. (referencing K. Sabeel Rahman & Jocelyn Simonson’s article \textit{The Institutional Design of Community Control}, observing that “some constituencies already possess greater capacity for power and influence than others,” 108 CAL. L. REV. 101 (2020)).
\textsuperscript{58} Id. at 774.
\textsuperscript{59} Id. at 774–775.
\textsuperscript{60} See generally, Abeba Birhane et al., \textit{Power to the People? Opportunities and Challenges for Participatory AI}, 2022 Equity & Access Algorithms, Mechanisms, & Optimism (promoting participatory approaches to AI).
\textsuperscript{61} Id.
2. What Kind(s) of Governance Does Blockchain Afford?

Blockchain governance on its functioning and underlying code, interactions among participants, and decision-making processes typically has two forms: on-chain and off-chain. On-chain governance is the governance rules that are encoded in the blockchain protocol and that govern interactions between participants. Since the rules are embedded within the code, this is sometimes referred to as the “rule of code.”

Off-chain governance refers to both internal and external governance that might impact the future development of the blockchain that is not embedded in the underlying blockchain protocol. This protocol includes both those rules adopted by a community of stakeholders and third-party decisions that impact the community, such as laws and regulations, contracts, and technical standards.

Blockchain governance may also entail governance that is distributed among stakeholders or a community of users through governance tokens. “Governance tokens give holders the right to participate in governance decisions . . . .” Scholars have argued that DLT may also promote democratic goals by lowering the social and economic barriers certain populations face especially by “bypass[ing certain] intermediaries and gatekeepers.” Novel institutional forms, such as DAOs, have arisen using blockchain technologies. DAOs enable governance and social coordination among actors based on a set of self-executing rules, or smart contracts, on a blockchain. DAO members hold tokens that permit participation in the organization, and decision-making is based on collective consensus mechanisms and voting. New governance institutions enabled by blockchain technology, such as DAOs, offer the potential for more “democratic or participatory decision-making.”

Blockchain technologies and polycentric governance can also help promote trust in public and private institutions and different models of

---

63. Id. at 822.
64. Reijers et al., supra note 62, at 822.
65. Id. at 822, 824.
68. Wright & De Filippi, supra note 26, at 3, 16.
social coordination at the global level. Blockchain technology offers a challenge to the current configurations of international legal regimes that centralize power, wealth, and governance. In fact, some have proffered that blockchain and the ideologies underlying the development of a trustless peer to peer DLT was a result of, and shaped by, the public’s loss of trust in both public and private institutions. Blockchain can also promote commons-based forms of governance in areas that are not limited to national jurisdiction, such as outer space.

There is an ongoing debate about which forms of governance, code-based rules or others, might work best to govern blockchain decision-making processes. Private powers pose challenges to code-based on-chain governance—which resembles formalist and positivist legality—just as they do to positivist legal systems which rely on liberal proceduralist conceptions of legitimacy. Moreover, since many blockchains are typically governed according to the number of tokens a participant has and not based on “one-person, one-vote,” proof of work based systems that rely on on-chain governance may lead to increasing corporate consolidation and plutocracy. Even though blockchain protocols offer distributed forms of governance, in emergency situations, the system may resort to core developers and decision-makers to decide on behalf of the community without democratic input. While blockchain technologies started in a decentralized or distributed manner, similar to the early internet, governance of blockchain networks is becoming increasingly centralized, and there is a need to counter market-based incentives that create antagonistic blockchain environments rather than collaborative ones. Power relations and dynamics within blockchains also pose challenges to democratic ideals of distributed governance.

70. Korhonen and Rantala, supra note 18, at 218.
73. Reijers et al., supra note 62.
74. Id.
75. Id. at 15–16.
76. Id. at 16–17.
78. Alston et al., supra note 14, at 721.
The potential for undemocratic private governance in blockchain technologies requires also examining public governance of blockchain technologies. Blockchain does not evade interaction with formal governance, regulation, and public institutions. One of the important questions is “how constitutional democracies can deal with the rise of private determinations shaping fundamental rights and democracy in the blockchain environment.”\(^79\) This requires consideration of how democratic values can be protected in the face of blockchain technologies,\(^80\) as well as whether these technologies hold democratic potential.

According to Korhonen and Rantala, blockchain is “essentially a political, not a technological, idea.”\(^81\) Yet its political orientation and how it is used depends on the communities that shape and use the technology and should not be reductively described as deriving from a solely libertarian ethos.\(^82\) Moreover, blockchain technologies are creating new centers and peripheries, especially with regard to the labor required to maintain the material infrastructures, the systems required for mining and maintaining DLTs, and the environmental impacts of these technologies.\(^83\) Proponents of peer-to-peer technologies have long claimed that these technologies hold the promise of evading censorship, monopoly, hierarchy, and centralized authority but have often overlooked their reliance on centralized computing infrastructures.\(^84\)

The self-regulatory mechanisms of blockchain technologies add an additional layer to the constitutional and juridical relations that are part of the assemblage of blockchain-based digital money. Many initially thought that blockchain would get rid of middlemen, including both public and private institutions, and that “blockchain technologies would promote the development of autonomous systems providing a parallel mode of governance competing with public authority, thus, building new private regulatory frameworks.”\(^85\) Yet the self-regulatory mechanisms of blockchain protocols that are built into the technical code and the

\(^79\). See generally BLOCKCHAIN AND PUBLIC LAW: GLOBAL CHALLENGES IN THE ERA OF DECENTRALISATION 6 (Oreste Pollicino & Giovanni De Gregorio eds., 2021).

\(^80\). Id. at 8.


\(^82\). Id. at 408.


\(^85\). Oreste Pollicino & Giovanni De Gregorio, Introduction in BLOCKCHAIN AND PUBLIC LAW, supra note 79 at 2.
consensus mechanisms designed for its governance also interact with formal legal and regulatory governance mechanisms, and these mechanisms are built upon an existing legal and institutional landscape. These interactions between *lex cryptographia* and state or other formal institutions must not be overlooked. The proliferation of (private) normative orders and legal fragmentation based on social sectors and their interdependencies give rise to new forms of conflicts of law, “intersystemic conflicts law” or “regime collisions.”

Modes of governance by blockchain technologies have also raised novel legal and jurisdictional questions. For example, Kavanagh and Ennis call blockchains “blockocracies,” or blockchain bureaucracies, due to their resemblance to traditional bureaucracies described by Weber. As they describe, “[i]n blockocracies, the official jurisdictional area is defined by the extent of the mining network, wherein the consensus protocol ensures that all miners share identical copies of the blockchain. Thus, a blockocracy is constituted by the network of these synchronized and Identical copies of the blockchain.” Moreover, as Goldenfien and Leiter describe with regard to smart contracts, or code that is capable of executing automated transactions, “[t]he combination of private agreements executed through computational mechanisms, effective dispute settlement, and the technologies that link them, give these legal and computational environments a ‘jurisdictional’ character.” These developments raise novel questions as to the extent these mechanisms of private ordering develop jurisdictional qualities and how they ought to interact with other legal and public institutions to protect the public interest.

3. *Democratic Money: Money as a Constitutional Project*

Traditional theories of money in neoclassical economics consider it a neutral medium of exchange. Yet critical legal scholars, such as Desan, have shown that money is political, legally configured, and impacts


89. Id. at 143.

how societies are ordered and governed. According to Desan, money “defines authority and distributes material as it operates.”  

In her theory, money, as a credit contract, can be designed in different ways to produce value and distribution in particular ways, often contributing to and helping produce inequalities. Those inequalities can be offset if money is designed to produce value differently. 

Redesigning money is essentially redesigning the credit contract, which itself has distributive and governance implications. Historically, sovereigns became the enforcers of monetary contracts, and underlying private law such as contract, property, and torts were obligations that sovereigns could enforce in monetary terms.  

Thus, systems of governance, ordering, and law were built around monetary contracts and enforcement. The British system, with the advent of the Bank of England, delegated its sovereign authority to issue money and control its supply to investors.  

This monetary system was premised on a Lockean individualistic profit maximization ideal of the economy, rather than a view of money as a public medium controlled by a sovereign for public ends.  

This ideal of the economy thereby became institutionalized through money. The political, economic, and social orders that were shaped by monetary design historically were therefore contingent on the interests that were embedded in the design choices. Monetary design choices can also “configure public authority and its relationship to members of the group, and the way people relate to each other.”  

Money has constitutional and normative qualities. It can shape institutional arrangements surrounding money’s materiality, money issuance, markets, and legal enforcement of credit contracts. Money is a constitutional project whose design entails “political determinations to represent value in particular ways.”  

Therefore, questions around money are ultimately questions of how to allocate power, governance,
and distribution in a political community and between public and private ordering. According to Desan,

> [M]oney has an internal design: societies produce it by structuring claims of value in ways that make those claims commensurable, transferable, and available for certain private as well as public uses. That architecture, in all its intricacy, determines the way money works in the world. Moreover, that architecture varies. As societies change the way they engineer money, they change its character and the market it makes.99

Money comes in the form of either tangible cash or currency, or bank money, which are credits to a creditor’s account with a bank, which is convertible to cash.100 Banks play a significant role in money creation through loans with clearing systems that allow settling debts against each other. Central bank deposits are essentially promises to pay money and back the convertibility of bank balances to cash. Commercial bank deposits with central banks are digital balances recorded on the central bank ledger.101

In our current system, private actors make political decisions as to how to allocate credit in society without any democratic authority.102 Digital money might offer novel ways to design and structure credit so that it is not solely distributed by private actors driven by profit and that it is more attuned to the needs of marginalized publics.

4. **Digital Money and Public Banking Institutions**

Digital money is also a governance project.103 Aims to democratize digital money originated in the desires to democratize networked communications technologies since the 1990s, with the interest by Cypherpunks and Cryptoanarchists in using cryptography to protect privacy and promote a free society where digital money could be used.

---

without government interference.\textsuperscript{104}

Contemporary debates on digital money have been flourishing around reimagining the role of public institutions, such as central banks, in light of current debates around CBDCs. Recent proposals have been made for both public digital cash and public banking services, including citizen accounts with central banks, such as FedAccounts\textsuperscript{105} or Citizen Central Banking,\textsuperscript{106} and postal banking for the unbanked or underbanked, to serve the interests of public citizens, allocate financial resources in a more democratic way, and restore trust in public institutions, such as the Federal Reserve.\textsuperscript{107} The merits of “Citizen Accounts” are that the Federal Reserve would be able to implement monetary policy directly with citizens through monetary deposits and interest payments rather than working through, and relying on, private banks.\textsuperscript{108}

These proposals are distinguishable from most discourses on CBDCs, which as Omarova notes, are “preoccupied mainly with operationalizing potential changes in central bank liabilities, rather than situating them within the broader institutional critique.”\textsuperscript{109} The CBDC discourses overlook and miss an important opportunity to rethink how the financial system operates.\textsuperscript{110} Instead, proposals for citizen accounts with central banks and public banking options through postal banking offer a reimagination of the current banking system and monetary policy. While these proposals may have predated the current discourse around cryptocurrencies, CBDCs, and e-cash, they have been revived in the current sociotechnical context.

\textbf{B. Global Dimensions of Digital Money and Monetary Sovereignty}

Digital money challenges traditional models of governance, sovereignty, and statehood, which requires novel and critical rethinking of these concepts as international lawyers and academics have come to understand them. Pistor argues, for instance, that “the emergence of

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{104} \textit{Lana Swartz, New Money: How Payment Became Social Media} 52–53 (2020).
\item \textsuperscript{107} See Crawford et al., supra note 106, at 116–21, 172.
\item \textsuperscript{108} Hockett, supra note 107, at 103.
\item \textsuperscript{109} Omarova, supra note 106, at 1235.
\item \textsuperscript{110} See id.
\end{itemize}
\end{footnotesize}
new forms of money and their governance structure will profoundly shape the nature of statehood to come.” Since the US dollar’s move away from the peg to gold in 1971, most fiat currencies are backed by the state rather than commodities. However, digital money poses a challenge to the global monetary architecture that has been in place since the 1970s, since it provides an alternative to state money, or in other words, “digital money that is backed not by the [state’s] power of coercion but by control of data.” In other words, statehood is increasingly linked to data.

While Pistor makes a compelling argument, especially as regards transformations in governance, monetary sovereignty, and statehood with digital money’s advent, she only considers privately issued money that a centralized entity controls—namely, the Libra currency (later called Diem) that Facebook (now Meta) proposed in 2019 and later scrapped. Digital money threatens monetary sovereignty, especially in the form of a private currency that has the potential to be used worldwide due to Meta’s global influence and network effects.

Her argument challenges traditional monetary-sovereignty conceptions and forms, especially considering the problems associated with any private, centrally controlled global digital currency becoming the dominant global currency. This dominant currency can be used for surveillance and data extraction to nudge, govern, and manipulate the global economy while having no financial-stability, equitable-distribution, or democratic-governance commitments. It can also contribute to the perpetuation of inequality and coercion, especially when with that power comes control of the underlying digital infrastructure.

The Financial Stability Board (FSB) has issued policy recommendations regarding “crypto assets” and stablecoins, especially due to the risk stablecoins pose to financial stability, as evidenced by the collapse of Terra. With the advent of blockchain technology, some cryptocurrencies have taken the form of stablecoins, which are issued

---

112. *Id.*
113. *Id.*
114. *See id.* at 8.
115. *See id.* at 8–9.
117. *See Lev Menand, Regulate Virtual Currencies as Currency*, Just Money (Feb. 14, 2020), https://justmoney.org/l-menand-regulate-virtual-currencies-as-currency/. Coins issued by private parties and pegged to fiat currency do not require a digital form or blockchain technology and have existed long before these technologies were first used. I
by private parties and pegged to fiat currency but not backed by governments. They promise that the holder can redeem them at face value, and banks and private parties can issue them on a blockchain. Yet recent events show that since governments do not backstop stablecoins, there can be liquidity runs, which result in holders not being able to redeem them for fiat currency. This was seen recently with the stablecoin, TerraUSD, which crashed in May 2022 and resulted in an approximately $600 billion loss. The risks that crypto assets pose, according to the FSB, are liquidity’s losses, leverage’s use, and crypto market’s interconnectedness, all of which can potentially impact the traditional global financial system. These risks are further amplified by the lack of transparency, inadequate consumer protection, and governance problems in this context.  

Three challenges the FSB has noted regarding regulation are the global actors involved, jurisdictional issues, and the opacity and distributed nature of a crypto asset network’s actors and governance, which makes accountability difficult to ascertain. Moreover, there is little agreement as to whether crypto assets are money, securities, or property, though the FSB proposes a functional approach whereby the “same activity, same risk, same regulation” framework can apply. This approach aims to extend traditional financial regulation to crypto assets depending on their functions.  

Digital money is shaping novel forms and variations in monetary sovereignty, as can be seen in the different approaches of the largest economies of the world, China and the United States. China has banned cryptocurrencies and has developed a CBDC, the digital renminbi (RMB), in an attempt to internationalize its currency and challenge the US dollar’s primacy in the international monetary system and the United States’ power over the Society for Worldwide Interbank Financial Telecommunication (SWIFT). Through shaping novel payment and monetary infrastructures using DLT and influencing international digital-currency standards, China may be able to protect its currency’s sovereignty and expand the international uses of its digital RMB, which until now has been mainly a domestic, territorial

Thank Daniel Rohde for indicating this. For the view of stablecoins as essentially synthetic money and their issuers as shadow banks, see id.  

119. See id. at 2–3.  
120. Id. at 4.  
121. See generally Huang & Mayer, supra note 7 (comparing policy approaches of China and the United States regarding monetary sovereignty, cryptocurrencies, and CDBCs).  
122. See id. at 329, 331.
currency. Digital currencies, according to Huang and Mayer, “can enable a transfer of monetary power, as the development of new monetary systems usually accompanies some dispersion of political power.”

Considering these transformations to the global monetary system, a benefit of taking a SC perspective is that it considers the multiplicity of normative orders simultaneously challenging (and reifying) monetary sovereignty.

III. DEMOCRATIZING DIGITAL MONEY’S DESIGN FOR THE PUBLIC GOOD

A. Embracing Pluralism and the Polycentric Governance of Digital Money

1. Public, Non-Blockchain Digital Money: DFC and the ECASH Act

While there have been proposals for using blockchain to protect privacy, none are as good as cash regarding protecting privacy from both public and private surveillance powers. Moreover, there has been some debate over the (in)compatibility of blockchain technologies and data-protection regulations, such as the European Union’s General Data Protection Regulation, and the difficulties in asserting jurisdiction over a blockchain network with worldwide nodes.

DFC, on the other hand, might hold more promise of democratic potential, as it can increase payment-services access for the unbanked and underbanked while upholding privacy in transactions, similar to

123. See id. at 335–39.
124. Id. at 335.
126. See generally Michèle Finck, Blockchains and Data Protection in the European Union, 4 EUR. DATA PROT. L. REV. 17 (2018) (examining and attempting to reconcile apparent differences between blockchain technology and the General Data Protection Regulation); Christopher Kuner et al., Blockchain Versus Data Protection, 8 INT’L DATA PRIV. L. 103 (2018) (challenging the view of Jan Albrecht, a Member of the European Parliament, regarding the General Data Protection Regulation’s probable incompatibility with blockchain); Christopher Millard, Blockchain and Law: Incompatible Codes?, 34 COMPUT. L. & SEC. REV. 843 (2018) (limiting Jan Albrecht’s view and noting that blockchain technology may be compatible with and have positive effects on the development of data protection law).
127. See BLOCKCHAIN AND PUBLIC LAW, supra note 79, at 8.
One of a DFC system’s benefits is the possibility to create along with it monetary, payment, and banking systems with greater payment-services accessibility for the public and the unbanked.

Following Grey’s proposed description, a DFC regime is a monetary regime with a universally accessible and interoperable payments network, managed by the government, which settles and stores balances in digital fiat currency units that are legal tender and are readily convertible at par with other forms of government currency. Such a regime could be implemented via universal retail central bank or public/postal bank deposit accounts and/or a decentralized e-wallet network in which wallet-holders exchange DFC balances without intermediaries via an open, secure transaction protocol, similar to certain kinds of mobile money transactions today.

Under such a regime, individuals could manage their own accounts or wallets, or, alternatively, could authorize payments intermediaries to make transactions from their wallet on their behalf. Intermediaries could even host or manage accounts or wallets on their clients’ behalf, similar to commercial email hosts or lawyers’ escrow accounts. Unlike with commercial banking, however, the individual would retain direct ownership over the DFC balance in question, rather than exchanging it for a generalized deposit liability against the intermediary.\(^\text{129}\)

A shift to a DFC regime might also allow broader public education and engagement on money, banking, and finance in ways that might not happen otherwise. The transition itself could be a democratic moment that “open[s] up new spaces for debate and political action in ways that improve the financial system.”\(^\text{130}\) Moreover, this shift would not be based on a blockchain and therefore would not require the energy-intensive Proof of Work blockchain-mining consensus mechanism, which


\(^{129}\) *Id.* at 174 (footnotes omitted).

\(^{130}\) *Id.* at 178.
has been shown to contribute to global carbon emissions at significant rates.\textsuperscript{131}

The US dollar's role in the global financial system as the dominant international-trade currency and as a reserve currency mean that implementing DFC policy may have ripple effects elsewhere, impact cross-border payment systems, and require overhaul of cross-border financial, banking, and monetary systems to integrate DFC. The US dollar's significant role in the international monetary system was notable after World War II, with the Bretton Woods system, the International Monetary Fund's (IMF) requirement that currencies and exchange rates be pegged to gold, and the US economy's strength in the postwar period.\textsuperscript{132} This significantly impacted global economic distribution, as the IMF placed "the burden of adjustment for balance of payment disequilibria . . . on debtor countries."\textsuperscript{133}

The US dollar's significance in the global monetary system increased after 1971, when Nixon did not convert US dollars into gold to avoid defaulting on the obligation to convert foreign dollars into gold and devalued the dollar.\textsuperscript{134} Other states followed, ending the pegged exchange-rates system in 1976.\textsuperscript{135} Liberalizing capital markets and tying US dollars to natural resources, like petroleum, ensured the US dollar could maintain its dominance as a reserve currency.\textsuperscript{136} Today, the US dollar has a significant role in transnational markets due to the rise of market-based finance.\textsuperscript{137}

In the United States, Rep. Stephen Lynch, Chair of the House Committee on Financial Services' Task Force on Finance Technology, introduced a bill in 2022 called the Electronic Currency and Security Hardware (ECASH) Act.\textsuperscript{138} The bill directs the Secretary of the Treasury to develop and pilot "an electronic version of the United States dollar . . . that replicates and preserves the privacy, anonymity-respecting, and minimal transactional data-generating properties of physical currency instruments such as coins and notes to the greatest extent technically and practically possible."\textsuperscript{139} It also establishes a Monetary Privacy Board to oversee the government's monetary-privacy

\textsuperscript{131} Bietti & Vatanparast, supra note 83, at 1–3.
\textsuperscript{132} See Feichtner, supra note 19, at 164–65.
\textsuperscript{133} Id. at 165.
\textsuperscript{134} See id. at 166.
\textsuperscript{135} Id.
\textsuperscript{136} Id. at 169–170.
\textsuperscript{137} BANK FOR INT'L SETTLEMENTS COMM. ON GLOBAL FINANCIAL SYS., US DOLLAR FUNDING: AN INTERNATIONAL PERSPECTIVE 22 (2020).
\textsuperscript{138} See generally ECASH Act, H.R. 7231, 117th Cong. (2022) (proposing a digital U.S. dollar and corresponding regulatory mechanisms).
\textsuperscript{139} Id. § 3(a).
E-cash under the ECASH Act would be anonymous, interoperable legal tender payable to the bearer. The government can issue E-cash on smart cards or mobile phones, even without internet access. The digital cash can be transferred to others without any fees or any recorded transaction ledgers. Digital cash offers the benefits of anonymity and privacy similar to physical cash but provides additional benefits in accessibility for the unbanked and underbanked, communities that have historically been excluded from financial institutions, and low-income communities with limited internet access, without having to pay transaction and exchange fees.

DFC has more democratic potential when combined with novel institutional arrangements in public banking, as previously discussed. It also has more democratic potential than CBDCs, which are not public oriented and would merely shift how central banks operate, account for their assets and liabilities, and interact with other central banks.

Designing DFC with privacy-preserving features and inclusion in mind can ensure that the public’s privacy interests in a democratic society are protected. According to Rohan Grey, digital cash initiatives plus transformations in the role of public institutions involved in monetary governance and financial access, such as central banks and postal banking, can increase accessibility and address public-oriented issues better than the current monetary and financial system in which profit-oriented private banks and fintech platforms have a strong role. Grey also advocates for monetary pluralism and experimentation and notes the benefits of taking a multi-institutional perspective to digital-money issues. As Grey points out, privacy-respecting digital money’s importance cannot be overstated: “Transactional anonymity, like anonymity more broadly, must be understood as a public good and a core bedrock of political freedom in a democratic society.”

While the DFC proposal soundly promotes democratic values and financial inclusion, it does not necessarily account for the needs of communities who lack access to money and whose political and economic subjectivity is not recognized, as this paper will discuss next.

140. Id. § 6.
141. See id. § 3(b).
142. Id. §3(c)(2) - (4).
144. Id. at 6–8.
145. Id. at 11.
2. Blockchain Currency for the Stateless: R-Coin

Both stateless and refugee populations are increasing due to conflict and are expected to increase significantly in the coming decades due to climate change. These populations pose particular challenges because they are often unbanked and typically can neither access nor resort to public or private institutions for financial and economic activity.

International organizations involved in giving humanitarian aid to these populations are increasingly using digital payments, digital cash, as well as cash aid using blockchain technology. However, these tend to be more centralized forms, such as a permissioned blockchain that is controlled by an international organization like the United Nations High Commissioner for Refugees (UNHCR). This control poses privacy concerns, as is the case whenever there is more centralized control over a blockchain and the data that is associated with transactions. The grant of cash aid using blockchain technology is often tied with biometric identity programs and identity projects on the blockchain, especially as stateless and refugee populations often lack identification documents.

There are also more bottom-up cryptocurrencies that are being developed. One example of a bottom-up approach to digital money is the R-Coin developed by a Rohingya and refugee led nonprofit for the largest stateless population in the world, the Rohingya, who are displaced throughout Malaysia, India, Bangladesh, and other countries. The R-Coin is a crypto token developed on a blockchain, designed to promote and incentivize community service and mentoring by stateless Rohingya and other refugees in Malaysia.146 As Malaysia is not a party to the 1951 Refugee Convention, the lack of administrative and public services for refugee populations in Malaysia means that these populations lack access to jobs, education, banking, medical services, and other basic public services.147 Since the Rohingya are not recognized as nationals of any country, they lack identification documents that further prevent them from accessing public services anywhere. For many Rohingya refugees, they lack legal and economic subjectivity due to their status as stateless persons.148 Blockchain technologies and self-sovereign identity systems seem to grant them a visibility and

---

subjectivity that they would not be granted otherwise. In the words of a
Rohingya refugee, Anis, “[t]hough I exist in the world my existence [is]
not recognized by the world. I am the invisible man of the visible
world.”

According to Prasse-Freeman, in contrast to techno-
governance tools like biometric identity cards developed by the
UNHCR that “promise to improve refugee lives but succeed in more
efficiently controlling them,” the Rohingya Project has appropriated
these technologies and redeployed them in an effort to transform, rather
than manage, refugee lives. While this project is not without
criticisms, such as ambivalence by some refugees and failure to address
structural violence and genocide that this community has faced, this
example shows the transformative potential of blockchain technology
and its creative uses in grassroots communities to meet some of their
immediate needs and to promote collaboration and community service.

Blockchain technology has also been used to help promote the
values of dissident communities in the autonomous region of Rojava in
Syria. Amir Taaki, a blockchain developer working to help build
governance structures for the communities living in Rojava, sees
blockchain as a technology that can enable spaces “where marginalized
communities can operate outside of state control” and to promote
decentralization, equality, sustainability, and diversity.

Local currencies, like R-Coin, “have the potential to serve as vital
laboratories of monetary experimentation.” While often limited in
terms of scalability, this experimentation can inspire broader reform
efforts and reimagining of money tied to systemic and global issues that
remain attuned to the needs of local communities. It remains to be
seen whether these projects will be supported by, or will inspire, the
design of novel public institutional forms and global monetary reform
that can help communities retain their democratic values. One could say
that the constitutional element of these communities is creating a novel

152. Korhonen and Rantala, supra note 81, at 411.
155. Id. at 518.
"jurisdictional"156 space for each community, where the blockchain network based on the informal governance, rules, and values set forth by and for the community operates in tandem with a variety of other constitutional orders.

B. Digital Money for the Public Good—Which Public(s) and Whose Good(s)?

1. Issues of Inclusion/Exclusion: The Imagined “Crypto Publics”

One of the risks with digitalization is that it “condition[s] participation and redistribution in the global digital economy, helping to generate new forms of marginality, precarity and inequality as well as compounding (or sometimes surmounting) old forms.”157 There are significant limitations of any political system or modes of decision-making and deliberation that are dependent on other social and economic factors, such as access to digital technologies, access to the internet, and understanding of digital literacy, which is far from evenly distributed, locally and globally. This is true also in the context of digital money.

While permissionless blockchain technologies claim to be decentralized158 and therefore more democratic in theory, the communities or users of those forms of digital money are a very small group of a population, usually within a certain demographic or age range and with access to the internet. For all the claims of inclusiveness, it remains a potentiality, due to issues of digital knowledge, literacy, and access. Democratization is therefore important because digital money has the potential to create a second layer of exclusion to public goods globally. It is essential to consider the tensions these issues of exclusion create for democratization of and through digital technologies.

156. Goldenfein & Leiter, supra note 88.
158. Despite numerous claims that take for granted that blockchain is a decentralized technology, this claim has been refuted. While permissioned blockchains are more centralized by design, even permissionless blockchains are not as decentralized as many crypto enthusiasts make them out to be. The emphasis on decentralization overlooks the significant role of online intermediaries and can be misleading. There is also an important distinction of public, permissionless blockchains that have centralized governance structures but distributed nodes for other functions, such as validating transactions. See, e.g., Angela Walch, Deconstructing “Decentralization”: Exploring the Core Claim of Crypto Systems, in CRYPTO ASSETS: LEGAL AND MONETARY PERSPECTIVES (Brummer & Chris eds., 2019).
2. The Role of Experts

Debates around blockchain and digital money have become a domain for “epistemic communities”\textsuperscript{159} to contest the vocabularies, epistemologies, and proper governance within their own domains of expertise. Experts are doing “boundary work” to fit the technologies within their own areas of expertise, to claim that they have the right knowledge to proclaim proper policy and governance frameworks and thereby construct their own authority. But this boundary work often excludes the very imagined publics these epistemic communities claim to be speaking for and acting on behalf of. This boundary work helps shape “epistemic jurisdictions” in which some experts are able to produce “technical knowledge for a given political community, topical arena or geographical territory.”\textsuperscript{160} These epistemic jurisdictions are part of the pluralistic normative orders that are shaping digital money and are also deriving their authority from it.

The relevant epistemic communities in the context of digital money may well depend on which form of digital money one is discussing. In the context of digital money created by private actors on a blockchain, the epistemic community might be the developers and those participating in the governance and the operations of the network, including miners and oracles. It would also involve the state and formal law to the extent that the digital money is created under the umbrella of an incorporated entity. Other relevant epistemic communities include international organizations, central banks, fiscal and economic policymakers, and experts.

The prospects of technologies of automation and those relating to blockchain are often met with anticipatory governance projects, characterized by increased managerialism, by international organizations like the OECD.\textsuperscript{161} This managerialism is not only a form of technocratic governance that is at odds with democratic governance but can also perpetuate inequality when universally applied.\textsuperscript{162} This is especially so for nonmember states in the Global South, which are often the subject of policy trials and experimentation, but have little say in

\textsuperscript{159} Peter Haas, *Epistemic Communities*, in 2 THE OXFORD HANDBOOK OF INTERNATIONAL ENVIRONMENTAL LAW 698 (Rajamani & Peel eds., 2021).


\textsuperscript{162} Campbell-Verduyn & Hütten, supra note 161.
the OECD’s deliberations.\textsuperscript{163}

Moreover, in policy reports on CBDCs, such as in the United States, there is a common reference to the promotion of democratic values, including privacy protections.\textsuperscript{164} Yet there is little or no public participatory aspect to the development and design of CBDCs, which poses risks that they will just be a continuation of the traditional international financial system in a digital form and that they are being propelled by a sense of competition with other CBDCs and private stablecoins in the international monetary plane.

As Prieto Munoz has noted about DLTs,

> the highly technical nature of these economic transactions means that a new epistemic community composed of programmers and fintech entrepreneurs is being empowered to set the vocabulary and conceptual categories that are going to determine how this system will be regulated at the national, international, and transnational levels.\textsuperscript{165}

Expert knowledge has long played a significant role in shaping the global political and economic order.\textsuperscript{166} Its limitations ought to be taken into account when considering what prospects digital money might offer for democratization, when many of the conversations and decisions surrounding it seem to be in the language of expertise and even more so when the conversations are about digital money at the global scale.

IV. CONCLUSION

This paper has argued that designing digital money toward democratic ends means embracing monetary pluralism and polycentric governance. Digital money affords novel and diverse modes of social organization, governance, and distribution that ought to be embraced and experimented with. Pluralism of digital money may also provide an opportunity to embed values that might otherwise be overlooked or not valued in market societies and for democratic experimentation with

\textsuperscript{163} Id.

\textsuperscript{164} The US Department of Treasury, \textit{The Future of Money and Payments: Report Pursuant to Section 4(b) of Executive Order 14067}, 36 (2022).

\textsuperscript{165} Jose Gustavo Prieto Munoz, ‘Crypto-investment’ in International Economic Law: A First Sketch, 20 GLOB. JURIST, 8 (2020).

\textsuperscript{166} DAVID KENNEDY, \textit{A WORLD OF STRUGGLE: HOW POWER, LAW, AND EXPERTISE SHAPE GLOBAL POLITICAL ECONOMY} (2018).
in institutional and technological designs to effect distributive, political, and ecological transformations. This is not a call, however, for yet more cryptocurrencies used for speculative and profit-oriented purposes. It is calling for monetary pluralism of digital currencies that have “non-monetary surplus values,” in other words, that are not driven by profit and extractive motives but rather driven by democratic aims to serve diverse publics’ interests. The institutional context and architectures of digital money have a significant impact on its prospects for effectuating democratic aims and care must be taken to address issues of digital exclusion and governance by experts in this domain. Experimentation and pluralism in this context should embrace public and not-for-profit institutional architectures over profit-oriented ones, as well as cooperative and commons-based approaches to governance, social coordination, and distribution enabled by digital and blockchain technologies.


168. Teubner, supra note 46.

Net Neutrality: A Fundamental Right in the Digital Constitution?

CHRISTOPH B. GRABER

ABSTRACT

Has net neutrality been institutionalized as a fundamental right in the digital constitution? Empirical evidence shows that statutory embedding of net neutrality has recently been taking shape in certain jurisdictions. The unfolding of constitutional structures and processes—the next step to be expected according to the theory of societal constitutionalism—is still pending. What would the scope of protection of net neutrality be if institutionalized as a fundamental right? To answer this normative question, two theoretically challenging issues must be addressed as a preliminary: First, what should be the adequate conceptualization of the relationship between the social and the technological? The perspective adopted by this paper is a combination between science and technology studies scholarship and legal sociology literature. Second, as net neutrality involves relationships between private parties, the question is how fundamental rights should be conceived beyond state-centrism. Methodologically, the sociological reflection of fundamental rights as institutions of society will serve as a benchmark for evaluating future developments of constitutional legal doctrine.

Keywords: Net neutrality, legal sociology, STS, systems theory, societal constitutionalism

1. Professor of Law, University of Zurich. The author would like to thank Bettina Heintz, Regina Kiener, and Gunther Teubner for comments and discussion.

Indiana Journal of Global Legal Studies Vol. 30 #2 (Summer 2023)
© Indiana University Maurer School of Law
I. Introduction

In 2018, members of the social democratic faction of the Swiss Parliament launched a parliamentary initiative to amend the Swiss Constitution with a new norm protecting net neutrality as a fundamental right. In their substantiation, the politicians found fault with net neutrality not even being mentioned in the Swiss Telecommunications Act, although the principle, arguably, is considered “an unwritten law” in Switzerland. In their view, net neutrality “is a consequence of freedom of expression and information” and should therefore be treated as a “constitutionally protected fundamental right.” The initiative was later withdrawn as the Swiss Parliament had meanwhile amended the Telecommunications Act, introducing a new provision on open internet (and net neutrality) at the statutory level. Developments aiming at statutory embedding of net neutrality are going on in other jurisdictions. However, endeavors to institutionalize net neutrality as a fundamental right have so far not grown further than the Swiss initiative.

Of course, an impressive number of attempts to craft an “Internet Bill of Rights” have been made during the last twenty-five years. In addition to state agencies, sponsors include diverse civil society organizations, business corporations, multi-stakeholder dynamics, and public international institutions. However, most of these efforts aspire simply to extend the reach of existing constitutional guarantees of communicative freedom, privacy, non-discrimination, etc. to the internet realm; and only a few include new rights and freedoms that are specific to the internet, such as access to digital information, due process in online content moderation, and net neutrality. Above all, they do not sufficiently address the key conceptual challenges of fundamental rights on the internet.

---

3. AS 2020 6159; BBl 2017 6559, art. 12e (Switz.).
Fundamental rights are commonly understood as those constitutional norms whose purpose is to protect the individual against the power of the nation-state. However, the constitutional questions that we encounter on the internet cannot be captured with the triad of individual, power, and state. Because the internet is a global medium, these questions mostly deal with issues going beyond the boundaries of a nation-state territory and national jurisdiction. Also, they often relate to the actions of private corporations. Since classic constitutional theory limits the effects of constitutional rights to state action, they are ineffective in situations where the infringer is a private (natural or juridical) person. Furthermore, classic constitutionalism’s methodological individualism contradicts the many real issues on the internet that transcend the realm of individual autonomy and demonstrate that free communication requires more than individual liberties.

To challenge classic theory, this paper adopts a functional perspective on constitutions which is based on sociological systems theory and the theory of societal constitutionalism. Consequently, some institutional (discursive and technical) guarantees are postulated as preconditions for the effective use of communicative freedom on the internet. Regarding the discursive preconditions, the European Court of Human Rights (ECtHR) argues that the “tools” that one finds on the internet are essential “for the participation in activities and debates related to questions of politics or public interest.” Similarly, the United States Supreme Court states that social media represents “what for many are the principal sources for knowing current events” as well as “the most powerful mechanisms available to a private citizen to make his or her voice heard.” This case law signals major courts’ increased willingness to acknowledge the discursive preconditions of individual communication online.

Regarding the technical preconditions of internet free speech, current fundamental rights practice mostly overlooks the technology’s material opportunities and constraints. The internet is a technical
facility that pre-shapes individual communication. As the internet’s infrastructure is a precondition for free speech, it should be included in the scope of constitutional protection. What would be needed though is a fundamental right that—beyond protecting online communication as a discourse—provides safeguards for an open network infrastructure.

As a preliminary to conceptualizing network infrastructure within the fundamental rights theory, the instrumental view of technology, which is characteristic of jurists, needs to be overcome. The technologies of the internet are not neutral. Rather, encoding prescribes—in a law-like manner—what a user can or cannot do online. This paper will introduce the concept of “technical affordance” as a starting point for a reconstruction of the relationship between the social and the material. Technical affordances are opportunities or constraints of a technology that are co-shaped by “design constituencies” and “impact constituencies” within processes of material design and social interpretation.

Accordingly, the sphere of communication and the sphere of material things interact with each other without either being able to determine the other.

The internet as a medium of mass communication is a relatively young phenomenon. Therefore, normative expectations about net neutrality are far from being settled. Rather, they are constantly evolving in legislation, jurisdiction, and civil society, as the net neutrality debates of the past decade in the United States, the European Union, and elsewhere show. This paper argues that these debates provide empirical evidence for an incremental constitutionalization of net neutrality as a fundamental right, banning discriminatory traffic management measures and guaranteeing an open internet architecture.

II. THE CONCEPT OF “CONSTITUTION” IN THE DIGITAL SOCIETY

The above-mentioned attempts to craft “Internet Bills of Rights,” the Swiss initiative to embed net neutrality as a fundamental right in the Swiss Constitution, and the civil-society driven net neutrality...
debates in the United States and elsewhere demonstrate the necessity to reflect on the concept of a constitution in the digital society as a prelude to an assessment of a postulated constitutionalization of net neutrality.

A. Formal and Functional Perspectives

When jurists and political philosophers use the word constitution, they generally refer to a formalized document established by a nation-state. Since the revolutions in the United States and France at the end of the eighteenth century, the notion of “constitution” has referred to a written text. Those texts symbolized the revolutionary project’s distance from the old regime, constituted state power, and legally regulated the exercise of state power. Jurists have been using the concept as a reference to a formal written document providing legal norms of the highest hierarchical level with primacy over any other type of law.

After the Second World War, a transformation of statehood and the effects of globalization contributed to a gradual erosion of the formal and state-centered understanding of constitution. Ensuing debates about a possible conceptualization of constitutionalism beyond national borders can be seen as a response to the realization of formal constitutionalism’s creeping decay. Several jurists have argued in the literature for a perception of international law as a hierarchical system of different and interrelated components with a constitution—represented by the Charter of the United Nations—at its top. Similarly, constitutional semantics have also been applied to the


17. Id. at 3–4.

interpretation of processes related to the World Trade Organization\textsuperscript{19} and the establishment of the European Convention on Human Rights (ECHR),\textsuperscript{20} as well as the primary law of the European Union, including the Charter of Fundamental Rights.\textsuperscript{21} This literature has generally been trying to extrapolate formal constitutional thinking from the nation-state to the global or regional level.\textsuperscript{22}

Debates about net neutrality, provisionally defined as a blueprint for a set of normative expectations about an open internet infrastructure, refer to problems that cannot be captured by a formal understanding of constitution. Formalized constitutional documents do not provide guarantees for net neutrality, neither on the national, international, nor supranational level. Because of the internet’s pervasiveness and the leading role of transnational telecommunications corporations in the net neutrality context, a state-centered approach to constitutions and fundamental rights is bound to fail.

Rather than a formal approach, a functional one has been employed in the literature to study constitutionalization processes from a sociological perspective. Constitutional sociologists have been distinguishing constitutive and limitative functions of a constitution. Historically, those functions evolved over a long time span since the end of the Middle Ages, before a pattern of statehood became visible.\textsuperscript{23} In the constitution of a modern nation-state, the constitutive function is typically anchored in those provisions which establish and organize the legislative, executive, and judicial powers. The limitative functions can be recognized in a constitution’s fundamental rights catalogue and in the checks and balances that regulate the relationship between the three branches of government.

Within a framework of sociological systems theory, Luhmann has conceived law and politics as two separate systems, arguing that system differentiation is a precondition for increased cooperation.\textsuperscript{24} As law and


\textsuperscript{22} Gunther Teubner, The Project of Constitutional Sociology: Irritating Nation State Constitutionalism, 4 TRANSNAT’L LEGAL THEORY 44, 52 (2012).

\textsuperscript{23} Thornhill, supra note 14, at 80.

\textsuperscript{24} Niklas Luhmann, Ausdifferenzierung des Rechts: Beiträge zur Rechtsssoziologie und Rechtstheorie 156 (1981).
politics are operationally closed systems excluding any input/output relationship, the constitution is a key mechanism for structurally coupling the two.\textsuperscript{25} The constitution of a nation-state has thus a different meaning in both systems: “For the legal system it is a supreme statute, a basic law. For the political system, it is an instrument of politics.”\textsuperscript{26}

The national territory scope of such a conceptualization of a constitution is in potential tension with Luhmann’s thesis that we are living in a world society. Rather than defined within national borders, world society is characterized by a general replacement of territorial differentiation by globally expanding processes of functional differentiation.\textsuperscript{27} While most sub-systems of society, including the economy, science, art, education, mass media, etc., extend globally, Luhmann specifies that the political system stays centered in the nation-state. Regarding the law, territorial differentiation continues to be important, but the system has expanded its reach to also include dynamics of international, transnational, and global law.\textsuperscript{28} Obviously, this asymmetry between the political and legal systems and the other spheres of society is a challenge for constitutional analysis in a transnational context.

Building upon Luhmann’s concept of a world society and dealing with literature by Philip Selznick,\textsuperscript{29} David Sciulli,\textsuperscript{30} Chris Thornhill,\textsuperscript{31} and others, Gunther Teubner, over the last two decades, has been designing a theory of societal constitutionalism.\textsuperscript{32} His perspective is one of transnational constitutional pluralism conceptualizing constitutions beyond the territorial boundaries of the nation-state and beyond the

\textsuperscript{25} Niklas Luhmann, Law as a Social System 410 (2004).
\textsuperscript{26} Id.
\textsuperscript{27} See generally Niklas Luhmann, The World Society as a Social System, 8 INT. J. GEN. SYS. 131 (1982).
\textsuperscript{28} Niklas Luhmann, 1 Theory of Society 96 (Mieke Bal & Hent de Veries eds., 2012).
\textsuperscript{29} See generally Philip Selznick Et Al., Law, Society, and Industrial Justice 3 (1969).
realm of politics.\textsuperscript{33} Teubner observes the evolution of constitutional norms in a variety of transnational contexts while showing a particular interest in the analysis of constitutional processes in the digital environment\textsuperscript{34} and in the economic sphere.\textsuperscript{35} One of the main questions in Teubner’s research relates to the role and constitutional status of transnational corporations (TNCs), as some of them have become so powerful that they challenge the governance of nation-states in a number of respects.\textsuperscript{36} TNCs have constituted themselves as new actors in the world economy to detach from the influence of national legal orders.

Teubner’s emphasis on the role of TNCs within his theory of societal constitutionalism is of value for the present paper, as transnational telecommunications companies and other major providers of internet access have been creating hybrid worlds of governance that deeply impact communicative freedom on the internet. The next section inquires as to what extent tech corporations have really been able to detach themselves from government regulation.

\textbf{B. Transnational or National Standards in Internet Law?}

The famous thesis that the internet evades state regulation was most eloquently pronounced by John Perry Barlow at the World Economic Forum in Davos in 1996.\textsuperscript{37} It epitomized the dream of “digital libertarians” and dominated cyberspace-related policy debates until the beginning of the twenty-first century.\textsuperscript{38} Three characteristics of the internet were invoked to underpin the thesis: (1) the technology of the medium, (2) the geographic distribution of its users, and (3) the digital specificities of the distributed content.\textsuperscript{39} Guided by the conception of the internet as a space beyond national political and legal control, regulators in the United States and the European Union opted for an

\textsuperscript{33} Teubner, \textit{supra} note 22, at 45.

\textsuperscript{34} Teubner, \textit{supra} note 32; Teubner, \textit{supra} note 22.


\textsuperscript{39} Boyle, \textit{supra} note 38, at 178.
approach of self-policing and internet exceptionalism. This is most clearly encapsulated in the so-called “safe harbor” rules in the US Communications Decency Act (DCA) of 1996. Section 230 DCA states: “No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.” Accordingly, internet intermediaries hosting or redistributing content that is provided by third-party users (hereafter Internet Content and Service Providers, or ICSPs) have been exempted from the effects of a range of laws that would otherwise apply. In 2000, the EU adopted horizontal (related to all kinds of content issues) liability rules in the E-Commerce Directive, providing for a “notice and takedown” self-regulation mechanism for ICSPs in Art. 14. In October 2022, the novel Digital Services Act replaced the E-Commerce Directive without, however, removing the liability “safe harbors.”

While the self-regulation approach has been prospering on both sides of the Atlantic for content-related services, a countermovement began for services related to the access to and use of the internet infrastructure. Around 1999, academics started warning about potential risks to the open internet infrastructure, and in 2003 Tim Wu coined the term “net neutrality” as a political principle demanding

---

46. See Edwards, supra note 43, at 94–95 (describing various types of services provided by internet intermediaries using a certain terminology); Christopher Marsden, Transnational Internet Law, in THE OXFORD HANDBOOK OF TRANSNATIONAL LAW 418, 424 (Peer Zumbansen ed., 2021) (describing various types of services provided by internet intermediaries using a different terminology).
47. CHRISTOPHER MARSDEN, NETWORK NEUTRALITY: FROM POLICY TO LAW TO REGULATION 29 (2016).
that “broadband carriers should not discriminate in how they treat traffic on their broadband network.” Despite its initial vagueness and absent legal foundation, net neutrality rapidly became a key benchmark in political debates, civil society dynamics, and regulations of the Federal Communications Commission (FCC) related to the future of internet infrastructure. The FCC’s 2015 Open Internet Order through a re-classification of broadband internet access as a common carrier service under Title II of the US Communications Act—provided for a solid legal basis for strict rules on net neutrality. However, after the installation of the Trump administration, the new rules were repealed in December 2017 and the net neutrality project in the United States was back at square one. In the EU, the debate about regulation of net neutrality was initially inspired by the US example and—after a slow start—eventually gained momentum and proceeded straight to the adoption of the Open Internet Regulation, which is a directly applicable EU legislation on access to the open internet and net neutrality.

Backed by the Court of Justice of the European Union (CJEU), those rules are being effectively enforced in the EU Member states. In the Telenor v. Hungary decision of September 15, 2020, the CJEU delivered the first interpretation of the Open Internet Regulation. At issue was one of Telenor’s tariff options by which a subscriber in Hungary benefits from a “zero tariff” for the use of certain services and applications, whereas the use of “non-zero tariff” services and applications is blocked or slowed down. The court found that such a scheme involves a measure interfering with network traffic, which is in breach of Article 3(3) of the Open Internet Regulation because it is commercially motivated.

52. GERLACH, supra note 49, at 247–49.
56. Id. at 54.
of net neutrality rules in the Open Internet Regulation. In the Vodafone v Germany decision of September 2, 2021, the court defined zero rating more generally as a commercial practice:

whereby an internet access provider applies a “zero tariff,” or a tariff that is more advantageous, to all or part of the data traffic associated with an application or category of specific applications, offered by partners of that access provider. Those data are therefore not counted towards the data volume purchased as part of the basic package. Such an option, offered in the context of limited packages, thus allows internet access providers to increase the attractiveness of their offer.57

Through incentivizing the use of certain applications and services in a non-transparent manner, such a practice is significantly constraining subscribers in the exercise of their rights as end users of internet access services. The Court’s conclusion is that zero-rating practices violate net neutrality and are thus unlawful under the Open Internet Regulation.

The upshot of these developments is that the thesis of a transnationalization of law in cyberspace is only undoubted as far as content-related regulation is concerned. Here, we observe self-regulatory approaches structuring the most important markets, although the potential implications of the EU’s novel Digital Services Act are still hard to predict. Regarding the internet’s technological architecture, however, the net neutrality debates reveal an orientation towards co-regulation approaches, allowing private actors to operate within a legal framework that is defined by national regulators.58 Are the regulation of the open internet infrastructure at EU and member state levels and the emergence of related case law of the CJEU evidence for renationalization of net neutrality?

Mainly in the context of recent geopolitical readjustments in relation to China and Russia, the specter of a “splinternet” has been evoked in the literature.59 As the Cybersecurity Act of 2019, the adoption of the Digital Services Act and the Digital Markets Act, and

57. Vodafone v Germany, CJEU, Case C-854/19, September 2, 2021, at ¶ 15.
58. Marsden, supra note 46, at 432 (arguing that net neutrality is ‘the internet’s exception to transnational law).
the proposed Regulation of Artificial Intelligence demonstrate, the EU has opted to go ahead with stronger regulation of the internet, in relative independence of US strategies or multilateral perspectives. Milton Mueller argues that debates about a technical fragmentation of the internet are overblown and distract from the fact that certain governments are striving to “re-align control of communications with the jurisdictional boundaries of national states.”\(^{60}\) With regard to net neutrality, I would agree that the predominance of national regulatory approaches should not make us overlook that net neutrality has emerged as a shared goal of most industrialized states. The process confirms Luhmann’s diagnosis of an ambivalence between territorial and functional differentiation regarding the systems of politics and law in the world society. This is underpinned by Thornhill’s thesis that constitutionalization processes proceed from transnational patterns to national law—and not the reverse.\(^{61}\) As we will see below, the reality of net neutrality debates in many regions of the world, including the European Union, the United States, and India, is evidence supporting this thesis.

### III. The Sociotechnical in the Digital Constitution

Broadly speaking, net neutrality is about securing an open internet infrastructure. Conceptualizing net neutrality within a framework of fundamental rights is thus facing the problem that the constitutional doctrine is “technologically blind.” Legal practitioners in general hold an instrumental view of technology; they perceive technology “as a neutral tool without examining its broader social, cultural, and political impacts.”\(^{62}\) Treating technology as a “black box,” legal doctrines and courts are mostly unable to understand the complex interactions between law and technology in the digital environment.

In the social sciences, the discipline that has gained the deepest knowledge about the relationship between digital technologies and society is science and technology studies (STS). While there are different strands within STS, a common feature of this sub-discipline of sociology is the rejection of technological determinism. Rather than viewing technology as something evolving according to an internal logic and independent from social influence, society and technology appear to be essentially connected. Although the perspective is one of social constructivism, human intention is juxtaposed with the material

\(^{60}\) Mueller, supra note 59, at 17.


\(^{62}\) Cockfield and Pridmore, supra note 11, at 476.
characteristics of an artifact. Avoiding “the charges of both natural and social determinism that have featured in recent academic debates around the field of science and technology studies,” the term “co-production” has become established as an idiom describing the entanglement of the social and the material in the design and uptake of new technologies.

“Technical Affordance” is a key concept to understanding how this process of co-production unfolds in detail. The concept of “affordance” was originally coined by James Gibson, a perceptual psychologist, in 1979. Gibson argues that animals can receive information in their environment selectively. What information is selected depends on its relative importance for the animal’s survival. Affordances are, thus, information concerning the environment’s possibilities or invitations. For Gibson, the natural environment exists independently of a living system’s representations of it. Consequently, affordances of things are independent of the perceptual experience an observer may gain from them.

Loosely based on Gibson, this paper uses the concept of (technical) affordance to describe the opportunities or constraints of technology. But how do affordances come into an artifact? The answer: in a process of co-production involving multi-layered interactions between the social and the material. Bryan Pfaffenberger introduced the “technological drama” as a heuristic to conceptualize this as a process of three acts. The drama starts in Act 1 with a design constituency shaping the material design of an artifact (“regularization”), followed by the response of an impact constituency in Act 2, which advocates an alternative interpretation of the technology’s ontology and purpose (“adjustment”). Act 3, “reconstitution,” is about a material redesign of the technology by the impact constituency. The heuristic illustrates

63. Helen Nissenbaum, From Preemption to Circumvention: If Technology Regulates, Why Do We Need Regulation (and Vice Versa)?, 26 BERKELEY TECH. L.J. 1367, 1378 (2011).
69. FEENBERG, supra note 65, at 153–57.
71. See Graber, supra note 9, at 199–206.
possible action patterns in the co-production of artifacts, comprising both discursive and non-discursive elements, and explains how affordances enter a technology.

Such an approach enables a reconstruction of the internet’s materiality within a perspective of legal sociology inspired by systems theory. The internet is conceived as a “structural coupling”\textsuperscript{72} of digital technology and communication, as a relationship of “dependent independence.”\textsuperscript{73} Hence, the internet is not a social system\textsuperscript{74} and it does not directly produce meaning. Rather, through structural coupling with social systems, the internet is interrelated with the level of complexity of society’s current form of differentiation. In the age of digitalization and AI-driven communication, complexity growth is spurred by machines that may attain a certain degree of autonomy while the affordances of “pre-emptive computing” will often remain hidden.\textsuperscript{75} This raises entirely new questions regarding legal normativity and the law’s ability to defend its autonomy against the threat of “algorithmic governmentality.”\textsuperscript{76}

IV. NORMATIVE EXPECTATIONS AND DIGITAL TECHNOLOGIES

From a sociological perspective, fundamental rights are bundles of normative expectations about the protection of individual and social autonomies against society’s self-destructing tendencies.\textsuperscript{77} Conceiving the design and uptake of a digital technology as a process of co-production of affordances creates scope for theorizing the emergence of normative expectations about a novel technology from the middle of society. Regarding net neutrality as a fundamental right, it is therefore necessary to reflect, as a next step, how normative expectations about technologies emerge.

\textsuperscript{72} See generally Luhmann, supra note 28, at 49–68, 322 (specifically on the structural coupling of the physical world and society).

\textsuperscript{73} Dan Wielsch, Zugangsregeln, Die Rechtsverfassung der Wissensteilung 236–38 (2008).

\textsuperscript{74} Contra Peter Bøgh Anderson, WWW as Self-organizing System, 5 Cybernetics & Human Knowing 5 (1998).

\textsuperscript{75} Hildebrandt, supra note 9, at 170. See also Bucher and Helmond, supra note 68, at 237; William W. Gaver, Technology Affordances, in Proceedings of the ACM CHI 91 Human Factors in Computing Systems Conference 79, 80 (Scott P. Robertson et al. eds., 1991).


\textsuperscript{77} See Graber, supra note 10, at 238; see also supra Section 5.1.
A. Emergence of Normative Expectations

From the teachings of Eugen Ehrlich, one of the pioneers of legal sociology, we know that a social institutionalization of norms predates, and is relatively independent of, their legal institutionalization.\textsuperscript{78} Besides Ehrlich, other classics of legal sociology have also emphasized the bottom-up emergence of norms. Emile Durkheim, for example, emphasizes the non-contractual foundations of the private contract,\textsuperscript{79} referring to binding behavioral expectations that may oblige parties, even in the absence of a formal agreement.\textsuperscript{80} Max Weber is similarly convinced that legal rules gradually evolve from behavioral expectations about social facts. He argues that the state-centered origin of law is the exception rather than the rule.\textsuperscript{81} Norms often originate “aboriginally,” starting with an individual’s subjective attitudes toward others, then evolve from mere habituation to awareness of others’ behavior, and eventually turn into contrafactually confirmed expectations.\textsuperscript{82} For Weber, an order of contrafactual expectations is a “convention” if the sanction for its violation is disapproval within a given social group. Only then can a normative order be called law, when there are external guarantees that it will be enforced with physical or psychological coercion.\textsuperscript{83} The source of such coercive power can be the state or other forms of legitimate social pressure.

Weber argues that normative patterns continue to emerge directly from social life “even in economically or socially complex and advanced societies.”\textsuperscript{84} While Weber’s object of observation was the relatively homogenous society at the turn of the twentieth century, today’s society is hypercomplex, heterogeneous, and highly fragmented. For Luhmann, the management of complexity and contingency in communicative processes is a central problem of modern society. While in relation to nature it is sufficient to learn through adaptation of behavioral expectations, adaptation among humans depends on learned expectations of expectations.\textsuperscript{85} For Luhmann, expectations are the

\begin{itemize}
\item \textsuperscript{78} Eugen Ehrlich, \textit{The Sociology of Law}, 36 Harv. L. Rev. 130, 138–42 (1922).
\item \textsuperscript{79} Emile Durkheim, \textit{The Division of Labor in Society} 165 (1997).
\item \textsuperscript{80} Id. at 170.
\item \textsuperscript{81} Max Weber, \textit{Economy and Society: An Outline of Interpretive Sociology} 753 (1978).
\item \textsuperscript{82} Id. at 754.
\item \textsuperscript{83} Id. at 34.
\item \textsuperscript{84} Id. at 753.
\end{itemize}
structures of a system of communication. He understands the building of cognitive and normative expectations as an evolutionary response to overwhelming complexity and contingency in communication. While cognitive expectations are adapted in case they are disappointed, normative expectations are generally upheld even if violations occur in singular cases. Even if expectations are held normatively, one can expect “that, in a cognitive context, normative expectations can also be changed through learning.” The law is the social system that is specialized in the stabilization and generalization of normative expectations. Because generally shared values are precarious in a hypercomplex society, the stabilization and generalization of normative expectations is demanding. Positive law has been the evolutionary solution to this challenge, allowing for flexible adaptation of legal norms to social change and providing, at the same time, for mechanisms securing legal certainty. However, only a small part of all expectations is institutionalized in the form of positive law, and, in everyday life, expectations continue to grow from the middle of society.

Regarding newly introduced digital technologies, a user’s expectations will depend on the technology’s expected behavioral “autonomy.” In relation to “predictable” artifacts, the adaptation of behavioral expectations will be subject-object related as in relation to nature. In relation to “smart” technologies (e.g., AI and machine learning algorithms) however, uncertainty will be an issue due to the expected unpredictability of the digital agent’s behavior. For Beckers and Teubner, personification would be the most successful strategy for coping with uncertainty in such situations, since personification, arguably, is transforming the relationship between a human and an algorithm “from a subject-object relation into an Ego–Alter relation.” According to Luhmann, personification is an evolutionary product of the law, primarily manifested in the concept of subjective rights and in the fiction of the legal person. Yet, Beckers and Teubner use

86. ALFONS BORA, Medialität des Wissens, in HERRSCHAFT UND WISSEN 295, 300 (Gunnar F. Schuppert et al. eds., 2022).
87. Luhmann, supra note 85, at 36.
88. See id. at 35–37.
89. Luhmann, supra note 25, at 108.
90. See id. at 147–48.
91. Id. at 261.
92. Luhmann, supra note 85, at 47.
93. Id. at 47–48.
95. See LUHMANN, supra note 25, at 269–71.
personification in a generalized way, without considering the difference between legal and non-legal communication. Legal communication presupposes a conflict because only where there is a conflict can a distinction between legal and illegal be operated at a level of second-order observation. Conflicts, however, are generally not an issue in day-to-day interactions between an ordinary user and an algorithm. Moreover, Beckers’s and Teubner’s suggestion to personify a human’s expectations of an algorithm’s expectations is in contradiction with their assertion (one page later) that “the algorithm’s internal operations cannot in any way be equated with the mental operations of humans.” Hence, personification does not strike me as a successful strategy to make uncertainty in non-legal user-algorithm interactions disappear.

If uncertainty remains the decisive issue in a relationship between a human and a digital agent, the question is how affordances can be produced at all. As a foundation to an answer, a distinction between ordinary users and expert users must be drawn. Ordinary users are in a synchronous relationship with their computer, which is recursive and updated at every click. Ordinary users do not understand how the hardware or software functions. They only see the affordances at the surface of the computer screen; the operations of the algorithm remain hidden in the background and only experts will be able to interpret them.

The possibility of an impact constituency to reject a design constituency’s interpretation of a smart technology’s affordances will largely depend on expert knowledge on the impact constituency’s part. Only experts and semi-professional activists may be able to detect the negative side effects of a smart technology, raise public awareness about its social risks, and mobilize civil society representatives and a broader public to build and invigorate alternative interpretations of the technology. If such a campaign is successful, normative expectations

96. See Beckers and Teubner, supra note 94, at 27.
97. See Luhrmann, supra note 25, at 245–46.
98. See Beckers and Teubner, supra note 94, at 27.
99. Id. at 28.
100. See Bettina Heintz, Big Observation: Ein Vergleich moderner Beobachtungsformate am Beispiel von amtlicher Statistik und Recommendersystemen, 73 Kölner Zeitschrift für Soziologie und Sozialpsychologie 137, 141 (2021) (describing distinction of the computer screen’s foreground and the “algorithmic backstage”); Beckers and Teubner, supra note 94, at 28–29 (similarly distinguishing between the screen “surface” and the computer “subface”, i.e., the inner workings of hardware and software, which remain hidden to the ordinary user).
about the artifact's affordances may eventually emerge. Legislation and judicial adjudication can both play a role in the decision on which normative expectations will be stabilized and generalized. At an early stage of a technology's uptake, normative expectations are mostly "unsettled." Accordingly, it will often be the courts who will first be in charge of applying the law in the context of the new artifact. At this point, they lack sufficient information about the socio-political impact of the technology. Legislation mostly sets in at a more advanced stage of settlement when the relevant political actors decide to draft a general and abstract framework regulating the new technology.

B. The Network’s Hidden Affordances and Strategies to Make them Visible

As indicated, the growth of normative expectations out of a process of co-production will end up with the problem of undetected or hidden affordances, that is, features of a technology that are not directly revealed to the ordinary user. In the case of net neutrality, extremely complex hardware and software is involved. How should an impact constituency be able to counteract a telecom corporation’s mythos about network traffic management or a tariff option and how should it be able to reconstitute the technology’s materiality despite asymmetries of power and knowledge in the relationship between the impact constituency and the design constituency?

The above-discussed decisions of the CJEU reveal that zero rating has been the main concern in cases involving net neutrality violation claims. Zero rating showcases ordinary users’ difficulties in discovering the technology’s affordances. The key problem is lack of transparency. Zero tariff options are often misleadingly marketed by telecom corporations, for example by making users erroneously believe that they receive access to the full internet when only a handpicked number of services can be accessed. A telling example is Free Basics (formerly Internet.org), a service that Meta (formerly Facebook) launched in cooperation with telecom companies in 2015. Free Basics was communicated in India and other developing countries as a philanthropic enterprise advancing internet connectivity in the global

101. See generally Graber, supra note 12 (discussing in detail the co-production of affordances and the growth of norms regarding Facebook’s like button).
103. See FEENBERG, supra note 65, at 32; Hildebrandt, supra note 9, at 170.
Research by civil society activists revealed that, in truth, Free Basics offered Indian users access only to a “walled garden,” a select number of websites.

The case underscores the important role of civil society experts in making the affordances of a complex technology transparent and triggering a regulatory response. In India, for example, the launch and marketing of Free Basics coincided with a public consultation by the Telecom Regulatory Authority of India (TRAI) on net neutrality. Recognizing the importance of this issue in relation to Free Basics, a group of activists initiated the “Save the Internet” campaign and mobilized hundreds of thousands of people to participate in the consultation. In February 2016, following the consultation, TRAI barred telecom service providers from charging differential rates for data services, thereby effectively banning Meta’s Free Basics. The Indian population seemed divided on the issue. Many vocal opponents of the protest criticized TRAI’s efforts, which they considered inhibiting a project potentially increasing internet access in rural areas. Yet, considering the overwhelming support that the net neutrality advocates received, the debate is deemed a milestone in demonstrating the effectiveness of internet activism in shaping public policy in India. Strikingly, it was not an institutionalized representative body but a few activist groups that raised awareness and advanced public mobilization in favor of net neutrality. While in 2015 TRAI still appeared minded to permit differential pricing, it was the pressure of public opinion and civil society activism that eventually led to the regulatory response.

106. Anita Gurumurthy and Nandini Chami, Internet Governance as ‘Ideology in Practice’ – India’s ‘Free Basics’ Controversy, 5 INTERNET POLICY REV. 1, 2 (2016) (also discussing the TRAI Consultation Paper on Regulatory Framework for Over-the-top (OTT) services that was launched on March 27, 2015).
108. Subhayan Mukerjee, Net Neutrality, Facebook, and India’s Battle to #SaveTheInternet, 1 COMMUN. PUB. 356, 360 (2016).
109. Id. at 360.
110. Esha Sen Madhavan, Internet and Social Media’s Social Movements Leading to New Forms of Governance and Policymaking Cases from India, 4 GLOCALISM no. 1, 21–22 (2016).
The question is how it would be possible to develop inclusive political procedures that would contribute to making affordances of tariff options and traffic management measures transparent. The Indian case demonstrates the crucial role of activist experts in detecting net neutrality abuses, something impossible for regular consumers.\(^\text{112}\) The measurement of net neutrality in a concrete case requires detailed knowledge about the technological and economic implications of internet access providers’ practices. Consumers depend on expert advice from internet activists to understand what is going on. Triggered by social pressure, internet regulators may commission specialist reports to investigate cases that have been flagged by activists as net neutrality violations.\(^\text{113}\) A further point relates to stakeholder engagement. The debates in the United States and the European Union reveal that imaginative campaigns have been successful in mobilizing citizens to stand up for their rights.\(^\text{114}\) To orchestrate such campaigns, websites and civil society groups need sufficient resources.

V. Net Neutrality and Fundamental Rights Institutionalization

As mentioned in the introduction, net neutrality has so far not been formally included in a constitution of a nation-state. Can we nonetheless observe an ongoing institutionalization of net neutrality as a fundamental right, as the thesis of societal constitutionalism holds?

A. Normative Expectations and Fundamental Rights as Institutions of Society

From a sociological perspective, Luhmann establishes a link between normative expectations and fundamental rights by understanding fundamental rights as institutions of society. According to Luhmann, an institution is a “complex of factual behavioral expectations, which are actualized in the context of a social role, and which can always count on social consensus.”\(^\text{115}\) Fundamental rights as institutions of society stand for a generalization of normative expectations and lead to a broadening of factual consensus.\(^\text{116}\) Fulfilling

\(^{112}\) Marsden, supra note 47, at 3.

\(^{113}\) See id. at 210.

\(^{114}\) See Graber, supra note 13, at 539, 550.

\(^{115}\) Niklas Luhmann, Grundrechte als Institution: Ein Beitrag zur politischen Soziologie 12–13 (1965).

a limitative function, fundamental rights serve to protect modern society’s functional differentiation against its self-destructive tendencies.\textsuperscript{117}

To understand this point, it is necessary to consider that functional differentiation has historically led to the emergence of several autonomous spheres of meaning, such as the systems of the economy, politics, law, science, art, education, religion, and mass media. This process has been enabled by symbolically generalized communication media including money, power, truth, and law. These media are also known as success media because they make the success of communication more likely.\textsuperscript{118} In the words of Teubner,

\begin{quote}
Inconvenient or bothersome communicative offers become successful as soon as a specialized communication medium furthers drastically the chances for their acceptance: the seduction of money for handing over goods and services, power threats for obedience to commands, theoretical or empirical proof for implausible assertions, the pressure of credits and exams for accepting learning offers, and normative persuasion for legal judgments.\textsuperscript{119}
\end{quote}

The stronger a system’s communication medium, the faster the system’s growth.\textsuperscript{120} Luhmann observes that success media may incite excessive growth developments in all function systems,\textsuperscript{121} raising the sociological question of how much “monetization, juridification, scientification, politicization” is thus produced.\textsuperscript{122} Reformulated in a normative perspective, the question shifts to assessing the risk that success media would stir such excessive growth that one function system’s hypertrophy would negatively affect or even destroy modern society’s differentiation overall.\textsuperscript{123}

As one response to such risks, fundamental rights have emerged as

\begin{itemize}
\item \textsuperscript{118} LUHMANN, supra note 28, at 122.
\item \textsuperscript{120} See id. at 509.
\item \textsuperscript{121} LUHMANN, supra note 25, at 272.
\item \textsuperscript{122} NIKLAS LUHMANN, \textit{THEORY OF SOCIETY}, Vol. 2, at 95 (2013).
\item \textsuperscript{123} Teubner, \textit{supra} note 119, at 513–14.
\end{itemize}
counter-institutions of modern society. Their function consists of making sure that an overly successful communication medium is not endangering social differentiation.

B. Emergence of Normative Expectations about Net Neutrality

The theory of social constitutionalism is Teubner’s response to the realization that society lacks a “stop rule” which would prevent system hypertrophy. Therefore, it is the task of social constitutions “to produce limits to destructive growth tendencies.” Teubner distinguishes three steps in the process of societal constitutionalism, starting with the emergence of normative expectations from a reflexive process within a social system (step one), followed by a juridification of those newly generated normative expectations in a structural coupling with the law (step two), and, finally, their constitutionalization within a reflexive process of the law (step three).

Applying this approach to the empirical observation of constitutionalization processes in the context of net neutrality, I suggest selecting the economy as the system with the highest probability that we can witness the emergence of normative expectations about network infrastructure. The accuracy of this selection is confirmed by the net neutrality debates in the United States and the European Union. In the United States, major telecom companies have been among the main protagonists defending their commercial interests as business companies. They seek to preserve their right to discriminate between high-speed services of corporate content providers and ordinary users, charging the latter a lower price. In their Petition for Review against the FCC’s Title II Order of February 2015, for example, the major US telecom firms tried to prevent a limitation of their economic power that would have resulted from enforceable rights of net neutrality. In the EU, the net neutrality debate was inspired by the US example. When the Open Internet Regulation was adopted by the EU legislator on November 25, 2015, civil society representatives criticized these rules as being the result of a “corporate interests come first” reasoning, since

125. Teubner, supra note 119, at 514.
128. GERLACH, supra note 49, at xv.
they did not explicitly exclude fast lanes or zero rating. While the telecom companies have been defending their business interests, civil society activists have countered that equal access to the network should not be sacrificed to economic reasoning, since internet openness is a precondition of free speech.

The interactions between the proponents and opponents of strict net neutrality rules indicate the successful operation of a reflexive process out of which normative expectations about network infrastructure have eventually emerged. The debates in the United States and European Union demonstrate that the telecom companies (and their lobby groups) benefitted in this process from an excess of economic power and knowledge about the functioning and effects of the technology at issue. The asymmetric bargaining relationship complicated the task for the proponents of net neutrality. On the other hand, consumer organizations, internet activists, and other civil society segments standing for net neutrality have shown to be highly motivated and very creative in their campaigns. This has proved to be one of the main reasons why the net neutrality proponents have succeeded in the United States mobilizing millions in the fight for their cause. In the European Union, Save the Internet, a broad coalition of civil society exponents and internet activists, was a crucial actor in convincing the Body of European Regulators for Electronic Communications (BEREC) to draft guidelines about the interpretation of zero rating and contested provisions of the Open Internet Regulation. Save the Internet created a web page providing legal analysis of the various official documents and assisted citizens with their participation in the public consultation. Through a colorful campaign, they managed to attract considerable media attention, motivate thousands of websites to protest, and mobilize a sizeable number of citizens to directly communicate with the telecom regulators. BEREC explained that the received comments

130. TEUBNER, supra note 32, at 90–92.
133. See Thomas Lohninger, How We Saved the Internet, EPICENTER WORKS (Oct. 4, 2016), https://epicenter.works/content/how-we-saved-the-internet-in-europe.
provided the regulators with valuable feedback from stakeholders and increased transparency.\textsuperscript{135}

\textbf{C. Juridification of Net Neutrality}

According to Teubner’s theory, what one would expect as the next step of constitutionalization is the juridification of the normative expectations that—in the first step—emerged from the bottom of society. Juridification means the reformulation of social norms into the language of the law. As a direct flow of information from one system to another is not possible, structural coupling is the only possibility for resonance between systems. Regarding the United States, there is empirical evidence that a structural coupling between the economic system and the legal system led to the FCC’s Title II Order of February 2015. In a multi-year communication process primarily with the Court of Appeals of the DC Circuit, the telecom lobby, and civil society, the FCC had been able to sharpen the normative contours of net neutrality and eventually adopt rules that were clear, rigorous, and enforceable.\textsuperscript{136} These rules were upheld in the Court of Appeal’s decision of June 14, 2016.\textsuperscript{137} The Court rejected the numerous challenges to the FCC’s reclassification of broadband services as telecommunications services that were brought forward by major US telecom companies. The Court ruled that the FCC’s 2015 Open Internet Order had a legal foundation and that the promulgated net neutrality rules were valid. However, because of the FCC’s reappointment under President Trump, the Commission abolished net neutrality in December 2017 and restored the classification of broadband services as information services.\textsuperscript{138}

In the EU, besides the legislator, BEREC, and the CJEU have been the main protagonists in the juridification of net neutrality. Influenced by the many comments on its draft guidelines for the implementation of net neutrality, BEREC provided for a narrow interpretation of zero rating and specified that throttling and blocking of services is prohibited when done for commercial purposes.\textsuperscript{139} In line with these guidelines, the CJEU—in the above-discussed case law on the Open Internet Regulation—concluded to prohibit, as a rule, zero-rating practices.


\textsuperscript{136} Graber, supra note 13, at 547.

\textsuperscript{137} U.S. Telecom Ass’n v. Fed. Commc’ns Comm’n, 825 F.3d 674 (D.C. Cir. 2016).

\textsuperscript{138} Restoring Int. Freedom, 17 F.C.C. 312 (2017).

\textsuperscript{139} Graber, supra note 13, at 550–51.
Regarding the EU, one can thus conclude that the juridification of net neutrality has been fully accomplished.

**D. Net Neutrality, a Fundamental Right: What Normative Evidence do we have?**

From a legal perspective, a fundamental right is commonly defined as a norm, contained in a constitutional text, that provides for a subjective right that can be claimed before a court against interventions originating from state action. Such a formal-normative approach will inform the following analysis, assessing to what extent the value of net neutrality has already been constitutionalized as a fundamental right in Europe.

Article 10 ECHR guarantees freedom of expression and information as a fundamental right. According to the case law of the ECtHR, Article 10 ECHR covers not only the form and content of information, but also the technical means used to transmit such information. The court’s conviction that the technical means of communication are an important condition for the effective enjoyment of freedom of expression became first visible in Autronic v. Switzerland, a case involving satellite technology. In that decision from 1990, the ECtHR held that “[a]rticle 10 applies not only to the content of information but also to the means of transmission or reception since any restriction imposed on the means necessarily interferes with the right to receive and impart information.”

It has been argued in the literature that the Autronic-based case law of the ECtHR suggests the Court’s (implicit) recognition of a “free-standing transmission right.”

The ECtHR’s more recent case law reflects that with the establishment of the internet as the primary distribution medium the technical means of transmission and reception have become a key condition of free speech. This case law has also been guiding the highest court of the European Union. Discussing the legality of upload filters, the CJEU most recently specified that Article 10 ECHR does not prohibit prior restraints on a means of dissemination as such. However, the Court continued “such restraints pose such a risk to compliance with that fundamental right that a particularly tight legal framework is

---

Arguably, net neutrality is an important aspect of the novel “right to transmit,” protected under Article 10 ECHR. Although the ECtHR has so far not explicitly addressed issues of net neutrality in the context of Article 10, it may do so in the future. This is because of the Court’s understanding of the ECHR as a “living instrument, which must be interpreted in the light of present-day conditions.” The judge-made law, which is emerging from such a dynamic interpretation of the Convention, is constantly required to prove its legitimation. This can be done, inter alia, by one of the Council of Europe’s Committee of Ministers recommendations, issued on key issues of social or technological development. Indeed, the ECtHR has been using such recommendations in a number of judgments “for the purpose of interpretation of the Convention.” With regard to an evolutive interpretation of Article 10 ECHR in the context of the internet, the Court will, at the given moment, most probably seek guidance from Committee of Ministers Recommendation CM/Rec(2016)1 on net neutrality. The main target of that recommendation is interferences with internet traffic that “result in blocking, discrimination or prioritization of specific types of content, applications or services.” For the Committee of Ministers, such practices raise concerns regarding freedom of expression, the right to private life, and the protection of personal data. Creating a direct link between Article 10 ECHR and net neutrality, the Committee continues to emphasize that

148. COMM. OF MINISTERS, Recommendation CM/Rec2016)1 of the Committee of Ministers to member States on protecting and promoting the right to freedom of expression and the right to private life with regard to network neutrality, COUNCIL OF EUR. Jan. 13, 2016), https://search.coe.int/cm/Pages/result_details.aspx?ObjectID=09000016805e1e59.
149. Id. at ¶ 2.
[t]he principle of network neutrality underpins non-discriminatory treatment of Internet traffic and the users’ right to receive and impart information and to use services of their choice. It reinforces the full exercise and enjoyment of the right to freedom of expression because Article 10 of the Convention applies not only to the content of information but also to the means of its dissemination.\textsuperscript{150}

The Committee of Ministers is obviously cross-referring to the ECtHR’s reasoning in \textit{Autronic}. Since the ECtHR has not yet established a direct link with net neutrality in its case law on Article 10 ECHR, the scope of the “right to transmit” is not clear. Would that novel right only work for the benefit of end-users? What about an internet access provider invoking Article 10 ECHR for the purpose of a complaint against a government measure regulating net neutrality, such as the EU’s Open Internet Regulation?

The Charter of Fundamental Rights of the European Union (CFREU) provides for a right in Article 11 that corresponds with Article 10 ECHR.\textsuperscript{151} While the CFREU is part of the primary law, the EU enacted detailed rules on net neutrality in the above-discussed Open Internet Regulation, which is part of secondary legislation. When interpreting the Open Internet Regulation, the CJEU has so far not established a direct link between net neutrality and the fundamental rights guaranteed in Article 10 ECHR or Article 11 CFREU.

VI. **Anticipated Next Step in the Constitutionalization of Net Neutrality**

As the above normative analyses showed, net neutrality has so far not been formally recognized as a fundamental right. In this section, I would now like to juxtapose the formal legal perspective with a functional empirical perspective on the constitutionalization of net neutrality.

\textbf{A. From Soft Law to Hard Law, from the Periphery to the Centre}

The direct or indirect recognition of net neutrality as a fundamental right by a constitutional court, such as the ECtHR, would be the next

\textsuperscript{150} Id. at ¶ 4.

\textsuperscript{151} Woods, \textit{supra} note 141, at 333.
step to expect according to the theory of societal constitutionalism.\textsuperscript{152} Soft law instruments, such as the above-discussed Committee of Ministers Recommendation CM/Rec(2016)1 on net neutrality or the Declaration on Digital Rights and Principles, proposed by the European Commission in January 2022, will most likely be among the main sources informing this further step in the constitutionalization of net neutrality. Aiming to explain shared political intentions of the EU, the Declaration emphasizes the key importance of technology for European citizens in the protection and promotion of their fundamental rights.\textsuperscript{153} “Connectivity” appears prominently among the specific commitments serving to implement these goals. In this regard, the EU commits to “protecting a neutral and open Internet where content, services, and applications are not unjustifiably blocked or degraded.”\textsuperscript{154} With regard to “participation in the digital public space,” the EU Commission commits to “supporting the development and best use of digital technologies to stimulate citizen engagement and democratic participation” and to “continuing safeguarding fundamental rights online, notably the freedom of expression and information.”\textsuperscript{155} In a further chapter, the Declaration underscores that “everyone should have the right to the confidentiality of their communications and the information on their electronic devices, and no one shall be subjected to unlawful online surveillance or interception measures.”\textsuperscript{156}

From a legal perspective, soft law is non-binding law,\textsuperscript{157} although it may play an important role in legal practice, as it reflects a normative consensus of contracting parties. From a sociological perspective, the above-discussed soft law instruments of the Council of Europe and the EU result from a structural coupling of the systems of law and politics. The two instruments are thus emerging from the periphery of the legal system, contributing to supranational (EU Declaration) and international (Council of Europe Recommendation) law-making.\textsuperscript{158} In the periphery, the legal system can absorb \textit{vox populi} and social pressures that have been building up in various sub-systems of

\textsuperscript{152} TEUBNER, supra note 32, at 110–13.
\textsuperscript{154} EU Declaration, supra note 153, at 3.
\textsuperscript{155} Id. at 5.
\textsuperscript{156} Id.
A transformation from soft law to hard law, and thus a shift from the periphery to the center of the law would occur, when a court, such as the ECtHR or the CJEU, would base the interpretation of net neutrality in the context of Article 10 ECHR or Article 11 CFREU on one of those instruments.

B. Net Neutrality: Prospected Scope of Protection of a Future Fundamental Right

What might be the scope of protection of a fundamental right of net neutrality? If defined narrowly, the new right would mainly encompass (1) the prohibition of unreasonable blocking and throttling of content dissemination, and (2) the prohibition of zero rating, that is, misleading business practices applying a reduced tariff to certain internet traffic only to increase the attractiveness of their offer.

A broader interpretation of net neutrality in the light of freedom of expression would extend to issues addressed in the Council of Europe Recommendation and the EU Declaration, such as the relationship between a secure internet infrastructure and privacy, freedom of expression, and democracy. As a precondition for free speech, users’ normative expectations in the integrity of the network infrastructure must be protected. People will only then be able to effectively use their freedom of expression if they can expect that their online communications will not be intercepted by or manipulated through non-transparent interventions from governments or private parties. Regarding illicit government interventions, a decision of the German Constitutional Court of 2008 is a landmark. In that decision, the Court recognized the protection of people’s expectations in the confidentiality and integrity of information technology systems as a new, unwritten constitutional right. The Court held that people relying on information technology systems for their communication should be protected in their expectations of the technological integrity and confidentiality of those systems. While the German Constitutional Court’s ruling is restricted to state action, I would argue that the new fundamental right should also provide protection against private actors monitoring or even manipulating internet users’ online behavior. Legislation would be better suited to implementing this claim than a court ruling. Theoretically, such interpretation reflects a re-entry of the

159. LUHMANN, supra note 25, at 291–304.
161. Id. at ¶ 181, 203–207.
162. Graber, supra note 10, at 251.
distinction between a legal and a sociological perspective on fundamental rights into constitutional rights doctrine.

In conclusion, the scope of protection of freedom of expression should be extended to cover the technology-related preconditions of individual communication online. Focusing on the means of dissemination in the digital environment, net neutrality should be a self-standing right informing not only court practice, but also government administration and legislation in all matters involving the internet. Beyond an individualistic-negative interpretation, the norm should thus include an institutional-positive dimension, providing for an obligation of the state to take legislative steps which are necessary to ensure effective use of the freedom also in relation to private actors. The term institutional-positive dimension is a fusion of national\textsuperscript{163} and international\textsuperscript{164} fundamental rights doctrine and as such a product of transnational constitutionalism.\textsuperscript{165}

\textsuperscript{163.} See Christoph B. Graber, Kulturverfassung, 3 VERFASSUNGSRECHT 2309, 2316–17 (Oliver Diggelmann et al. eds., 2d ed. 2020) (exemplifying freedom of artistic expression and how the terminology varies depending on national jurisdiction. In Switzerland, the terms programmatic or institutional dimension of fundamental rights are used. The recognition of an institutional dimension considers that individual and social autonomies must be protected against any system with expansive tendencies).

\textsuperscript{164.} See KÄLIN & KÜNZLI, supra note 157, at 87–112 (explaining that in international law, human rights are conceived as obligations of states to respect (negative duty) as well as to protect and to fulfill specific rights through legislation (positive duties)).

\textsuperscript{165.} Thornhill, supra note 61, at 143 (using the term “transnational mutation” to refer to fusions of international law with national law).
Against Procedural Fetishism: A Call for a New Digital Constitution

MONIKA ZALNIERIUTE

ABSTRACT

Digital constitutionalism, to date, has been proceduralist; it has assumed that transparency and due process can temper power and attain justice for people vis-à-vis the automated state and powerful tech companies. So far, digital constitutionalism has also been very soft and blind to its own coloniality: Instead of deploying hard law, we are still looking for ways to pressure digital behemoths to self-regulate. We downplay US dominance, colonial exploitation, and environmental degradation caused by digital imperialism. Meanwhile, the power of tech companies has escalated. They now influence many aspects of our public and private lives, from elections to our own personalities and emotions, to environmental degradation. To be successful, the project of digital constitutionalism must resist a corporate agenda of procedural fetishism, a strategy to redirect the public from more substantive and fundamental questions about the concentration and limits of power to procedural microissues. Such diversion merely reinforces the status quo. To rectify the imbalance of power between people and tech companies, a new digital constitution must therefore try something different. It must shift its focus from soft law initiatives to tangible legal obligations by the tech companies. We must redistribute wealth and power not only by breaking and taxing tech companies, fortifying regulatory enforcement, increasing public scrutiny, and adopting prohibitive laws but also by democratizing big tech by making them public utilities and giving people a say how these companies should be governed. Crucially, we must also decolonize digital constitutionalism through recognition of colonial practices of extraction and exploitation and paying attention to the voices of

1. Senior Lecturer and ARC DECRA Fellow at the Faculty of Law & Justice at UNSW Sydney; Senior Fellow at the Law Institute of the Lithuanian Centre for Social Sciences. I thank Angelo Jr Golia, Gunther Teubner, Sofia Ranchordás, Tatiana Cutts and all participants of the Workshop on Digital Constitutionalism (15th June 2022) for invaluable feedback.

Indiana Journal of Global Legal Studies Vol. 30 #2 (Summer 2023) © Indiana University Maurer School of Law
Indigenous peoples and communities of the so-called Global South. With all these mutually reinforcing efforts, a new digital constitution will debunk the corporate and state agenda of procedural fetishism and will establish the new social contract for the digital age.

I. INTRODUCTION

The public eye is on tech companies aiding governments in cracking down on climate protesters;\(^2\) using the Covid-19 pandemic to leverage a greater role in education and state functions;\(^3\) profiteering through privacy exploitation; using psychological techniques to manipulate user emotions;\(^4\) and engaging in anticompetitive practices.\(^5\) These are just several examples of charges of abuse of power by tech companies, such as Facebook, Google, and Amazon.

We need a digital constitution to tackle these abuses of power. However, our efforts to date have been proceduralist. We have assumed that transparency and due process can temper power and protect the interests of people against digital behemoths and automated state power. Grounded in what I call procedural fetishism, our digital constitution has failed. We are still looking for ways to pressure tech companies to impose effective constraints on their escalating power to influence many crucial aspects of our public and private lives, from our elections to our personalities and emotions, to environmental degradation through the extraction of global resources to labor exploitation.

In this paper, I look at the laws, policies, and initiatives of the digital constitutionalism project at national, regional, international, and transnational levels and develop the concept of procedural fetishism. I show how the corporate (and state) agenda of procedural fetishism—an overemphasis and focus on procedural safeguards—acts as an obfuscation and redirection of the public from more substantive and fundamental questions about the concentration and limits of power by the tech companies to procedural microissues and safeguards. Such redirection merely reinforces the status quo. Procedural fetishism

---

3. *Id.*
detracts from the questions of substantial accountability and obligations by diverting attention to “fixing” procedural microissues that have little chance of changing the political or legal status quo. The concept of procedural fetishism enables us to see how digital constitutionalism has been completely blind to colonial extraction practices, labor exploitation, and dominance of the US tech companies, as if these did not exist. Procedural fetishism—whether corporate or state—is dangerous. Not only does it defer social and political change but it also legitimizes corporate and state power under an illusion of control and neutrality.

Proponents of proceduralist interventions in digital constitutionalism argue that these interventions are substantial steps toward justice and “an essential part of any comprehensive solution to regulate platforms.” True, interventions are essential. As Thomas Main argued, substantive law is neither procedural nor transprocedural. The point is not that procedure does not matter. I argue that procedures alone—devoid of substance—are dangerous for obfuscating the largest substantive problems we are facing today. We should resist procedures devoid of substance, which are often employed to distract from substantive problems, or legitimize corporate practice, business models. It is this kind of empty proceduralism, or what I call procedural fetishism, that is plaguing our digital constitution. Procedures need to come together with substantive policy and law. The latter has been lacking in digital constitution, and we must change that. This aspiration is consistent with the broader theoretical approach of societal constitutionalism in two senses. First, societal constitutionalism’s critical, transformative, and contestation potential has so far been unduly overlooked by digital constitutionalism scholarship and is characterized by a critique of the excessive reliance on procedures tailored to the traditional models of Western, liberal, state-centered constitutional law. Second, societal constitutionalism focuses on conflict-driven social dynamics and on the collision/disentanglement of distinct rationalities, notably the economic and the political ones, as a way to generate effective constitutional rules.

9. Id. at 360, 373–74, 390, 404–05, 407. For the deconstruction towards such models in the specific field of digital constitutionalism; see also Gunther Teubner & Angelo Jr Golia, Societal Constitutionalism in the Digital World: An Introduction, 30 IND. J. GLOBAL LEG. STUD. (2023).
10. Teubner & Golia, supra note 9.
In this sense, societal constitutionalism provides valuable, normative directions to rectify the imbalance of power between people and tech companies. For this purpose, this paper argues that the new digital constitution must shift its focus from soft law initiatives to substantive accountability and tangible legal obligations by tech companies. We must redistribute wealth and power not only by breaking and taxing tech companies, fortifying regulatory enforcement, and increasing public scrutiny by adopting prohibitive laws but also by democratizing big tech companies by making them public utilities and giving people a say on how these companies should be governed. Crucially, we must decolonize digital constitutionalism through recognizing the colonial practices of extraction and exploitation and listening to the voices of Indigenous peoples and communities of the so-called Global South. With these mutually reinforcing efforts, a new digital constitution will debunk the corporate and state agenda of procedural fetishism and will establish a new social contract for the digital age.

II. THE SOFT DIGITAL CONSTITUTION

Digital constitutionalism is about constraining power in the digital age. While the term has recently gained popularity among academic, civil society, and policy-making circles, it is not new. In various forms, constitutionalist discourse surrounding the need to tackle rising private power has been featured in different contexts over the last two decades. For example, in the 2000s, Paul Schiff Berman argued for “constitutive constitutionalism,” the idea that the “[c]onstitution might appropriately be viewed as a touchstone for articulating constitutive values and for structuring public debate about fundamental social and political issues,” which included the governance of digital platforms and what was then called “cyberspace.” In 2010, Nic Suzor coined the term “digital constitutionalism” we use today and described it as “the work of articulating limits on the exercise of power in a networked society.” Others defined digital constitutionalism as the “constellation of initiatives that have sought to articulate a set of political rights,

Against Procedural Fetishism

governance norms, and limitations on the exercise of power on the Internet."14 While scholars have differing views on the scope of digital constitutionalism and how to best implement it,15 all agree it should cover both public and private power. The traditional constitutional focus on public actors cannot account for the importance and dominance of private actors in digital society. We must therefore decouple “constitutionalism” from the state dimension.

Digital constitutionalist efforts in practice also cover a wide range of policies, laws, and voluntary initiatives at the national level, including domestic constitutions, laws, and judicial decisions; regional and international instruments and jurisprudence; self-regulatory initiatives; and transnational, nonbinding guidelines developed by private actors and NGOs.

Many recent digital constitutionalist efforts aim to tackle private tech power with national laws. For example, in the United States, five bipartisan bills collectively referred to as “A Stronger Online Economy: Opportunity, Innovation and Choice” have been proposed and seek to restrain tech companies’ power and monopolies.16 In addition, thirty-one state-level bills regarding digital privacy17 and eighteen for regulating social media platforms for taking down political content were introduced in 2021 in the United States.18 In China, the tech companies once seen as untouchable (particularly Alibaba and Tencent) have faced a tough year in 2021.19 For example, the State Administration for Market Regulation (SAMR) took aggressive steps to reign in monopolistic

15. See Celeste, supra note 8, at 87.
behavior, levying a record US$2.8 billion fine on Alibaba.\textsuperscript{20} Tech companies in Australia are also facing regulatory pressure targeting anticompetitive behavior.\textsuperscript{21}

Digital constitutionalism is also prominent at a regional level. Perhaps the strongest example is in the European Union, where several prominent legislative proposals have been tabled in recent years. The Digital Markets Act\textsuperscript{22} and Digital Services Act,\textsuperscript{23} as well as the Artificial Intelligence Act,\textsuperscript{24} and the Data Governance Act\textsuperscript{25} aim to tame the power of private corporations and limit the use of AI systems. These proposals build on the EU’s strong track record in digital constitutionalism. For example, the EU Directive on Copyright in the Digital Single Market\textsuperscript{26} and the EU General Data Protection Regulation\textsuperscript{27} have limited tech companies’ discretion in content moderation and the processing of personal data respectively. The EU has been leading digital constitutional efforts on a global scale with its binding laws and regulations. It has also attempted to fortify its leading positions with declaratory instruments, such as its recent declaration of rights and principles in the digital age, adopted by the EU Commission in 2022.\textsuperscript{28}

At an international level, digital constitutionalism has often resorted to human rights principles in attempts to tame corporate power. The Universal Declaration of Human Rights (UDHR)\textsuperscript{29} and the main instruments through which the UDHR has been codified, the

\begin{itemize}
\item \textsuperscript{20} \ Andrew Ross Sorkin et al., \textit{Alibaba’s Big Fine Is a Warning Shot}, \textsc{N.Y. Times}, (Apr. 12, 2021), https://www.nytimes.com/2021/04/12/business/dealbook/alibaba-fine-antitrust.html.
\item \textsuperscript{24} \ Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts COM (2021) 206 final (Apr. 21, 2021)
\item \textsuperscript{25} \ Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) COM (2020) 767 final (Nov. 25, 2020).
\item \textsuperscript{28} \ See Commission Proposal for a Declaration on Digital Rights and Principles for the Digital Decade, COM (2022) 28 final (Jan. 26, 2022).
\item \textsuperscript{29} \ See G.A. Res. 217 (III) A, Universal Declaration of Human Rights (Dec. 10, 1948).
\end{itemize}
International Covenant on Civil and Political Rights, and the International Covenant on Economic, Social and Cultural Rights, as well as fundamental rights set out in the International Labour Organization’s Declaration on Fundamental Principles and Rights at Work, have been main sources of inspiration. However, despite the popularity of the human rights discourse in digital constitutionalist projects, international human rights instruments are not directly binding on private companies. Instead, various networks and organizations try to promote human rights values among tech companies. For example, the Freedom Online Coalition (FOC) is a group of thirty-four governments, committed to work together to support internet freedom and protect fundamental human rights “through means of diplomatic coordination, shaping global norms, and multistakeholder collaboration.” Despite high-level governmental commitment to digital rights, many governments have not implemented enforcement mechanisms to ensure compliance.

Lacking enforceable constraints and obligations for tech companies in national or international law, digital constitutionalist efforts have therefore often been forced to rely on self-regulation through which tech companies assume responsibility for their own rules or practices and overseeing any sanctions for noncompliance. These range from various voluntary commitments to codes and soft law pronouncements to pseudocourts, such as Facebook’s Oversight Board. Self-regulatory instruments are often grounded in a concept of corporate and social responsibility (CSR). A good example of a global CSR instrument is the UN Guiding Principles on Business and Human Rights, unanimously

36. Zalnieriute, supra note 30, at 304.
endorsed by the UN Human Rights Council in 2011. The principles provide that governments retain the exclusive responsibility to protect and fulfill human rights obligations, while corporations are required to respect human rights. Corporations should avoid causing or contributing to human rights impacts, seek to mitigate them, and have policies and procedures in place, such as due diligence, to prevent and mitigate such impacts. Several other initiatives have developed voluntary principles for the tech companies. For example, the Global Network Initiative Principles on Freedom of Expression and Privacy, the Manila Principles of Intermediary Liability, the African Declaration on Internet Rights and Freedoms, and the Ranking Digital Rights Corporate Accountability Index seek to ensure that tech companies do not abuse their power to transparently communicate with their users and to remain accountable to the public.

A. The Limited Frame of Digital Constitutionalism

However, the digital constitution, to date, has had only limited success in taming the power of tech companies and dealing with global inequalities arising from that power. This weakness stems from the proceduralist focus of the digital constitutionalist project. Proponents have assumed that procedural safeguards, transparency, and due process can temper power and protect the interests of people against the power wielded by tech companies (and the state) in the digital age. Such assumptions stem from the normative foundations of digital constitutionalism and a popular analogy between tech companies and states: how platforms such as Twitter, Facebook, Google, and Amazon exert quasi-sovereign influence over commerce, speech and expression, elections, and other areas of life. It is also this analogy, and the power

38. Id. at 6, 19.
39. Id. at 19.
40. Id.
41. Id. at 20.
46. For literature making such analogies see generally Julie E. Cohen, Between Truth and Power: The Legal Constructions of Informational Capitalism (2019). See also Julie E.
of the state as the starting point, that lead to the proceduralist focus and emphasis in digital constitutionalism. Just as the due process and safeguards constrain the state, they must now also apply to tech companies.47

Yet, by focusing our energy on procedures, digital constitutionalism misses the larger picture. We rarely (if at all) discuss the US dominance or environmental exploitation and environmental degradation caused by digital technologies. Digital constitutionalists rarely ask how tech companies and their technologies reinforce existing power disparities globally between the so-called Global South and Imperialist West/North. Or how they contribute to climate disaster and exploitation of people and extraction of resources.48 These are the most important issues we face, yet we are often too busy designing the procedural, microlevel fixes.

To be successful, the project of digital constitutionalism must resist what I call a corporate agenda of procedural fetishism: a strategy, employed by tech companies and state actors to redirect the public from more substantive and fundamental questions about the concentration and limits of power in the society to procedural safeguards and microissues. This diversion reinforces the status quo, reinforces Western dominance, accelerates environmental degradation, and exploits the postcolonial peoples and resources.

III. PROCEDURAL FETISHISM

The term “procedural fetishism” was coined by William H. Simon in 1978 in his discussion of the ethics of the legal profession.49 Simon argued that the lawyers often believe in “an inherent value or legitimacy” of judicial proceedings.50 Such belief enables lawyers to

---

47. Danielle Keats Citron’s and Frank Pasquale’s early ground-breaking calls for technological due process has been influential, see generally Danielle Keats Citron, Technological Due Process, 85 WASH. U. L. REV. 1249 (2007) Although Citron’s original work did not focus on tech platforms, but argued that administrative agencies’ use of technology should be subjected to due process; also see generally Danielle Keats Citron & Frank Pasquale, The Scored Society: Due Process for Automated Predictions Essay, 89 WASH. L. REV. 1 (2014) (arguing for due process for automated credit scoring).

48. For a recent contribution to this debate, see generally Zahra Takhshid, Regulating Social Media in the Global South, 24 VAND. J. ENT. & TECH. L. 1 (2022).


50. Id. at 38.
justify their actions as ethical regardless of the consequences, merely because they are a part of this (supposedly) legitimate process.\(^{51}\) According to such view, a lawyer must follow his role “regardless of whether doing so will promote the discovery of truth.”\(^{52}\) The goal of the procedures themselves—the pursuit of truth—is what legitimizes decisions and renders them ethical.\(^{53}\) By fetishizing the legal procedure, the lawyers can rationalize practices that subvert substantive values such as autonomy, responsibility, and dignity.\(^{54}\)

The trend toward the proceduralism in modern law and politics and how it enhances the legitimacy of public (and private) authorities has been debated for at least four decades.\(^{55}\) Proceduralism, in its broadest sense, refers to “a belief in the value of explicit, formalized procedures that need to be followed closely,”\(^{56}\) or “the tendency to believe that procedure is centrally important.”\(^{57}\) The term is often used to describe the legitimization of rules, decisions, or institutions through the process used to create them, rather than by their substantive moral value.\(^{58}\) A rule becomes “acceptable” if it is “the outcome of an agreed-upon procedure.”\(^{59}\) Under such a view, violations of fundamental rights can also be legitimate, as long as they are authorized by the designated procedure.\(^{60}\)

Such a trend toward proceduralism, or what I call procedural fetishism, also dominates our digital constitution. We believe that having certain “safeguards” by tech companies (or governments) is inherently valuable, that those safeguards provide sufficient grounds to trust substantive corporate policy and practice, and that safeguards can, and often do, tame power.

\(51.\) Id.

\(52.\) Id.

\(53.\) Id.

\(54.\) Id. at 127.


\(60.\) Rocheleau, supra note 58.
A. Procedural Fetishism Enables the Appearance of Political Neutrality

Procedural fetishism is powerful and dangerous for several reasons. First, procedural fetishism offers an appearance of political and normative neutrality, which is convenient to tech companies, policymakers, judges, and regulators. Proceduralism allows various actors to “remain agnostic towards substantive political and moral values” when “[f]aced with the pluralism of contemporary societies.” At the heart of all proceduralist accounts of justice, therefore, is the idea that, as individual members of a pluralist system, we may agree on what amounts to a just procedure (if not a just outcome), and “if we manage to do so just procedures will yield just outcomes.” However, procedural fetishism enables various actors not only to remain agnostic but to avoid confrontation with hard political questions. Through prioritization of the manner and methods of pursuing substantive norms over specifying or imposing the substantive norms themselves, the end point of the process, as Mihalis Kritikos has noted in the context of EU law, with the procedure taking center-stage is “abandoned.” Similarly, the courts engage in procedural fetishism to appear neutral and to avoid tackling the politically difficult questions of necessity, proportionality, and legitimacy of corporate and state surveillance practices, and they have instead come up with procedural Band-aids that divert attention away from these more substantive and fundamental issues. The focus on procedural safeguards provides a convenient way to make an appearance of effort to regulate without prohibiting any practices or conduct. Proceduralist interventions and focus, instead of being neutral, can thereby reinforce the existing institutions and power disparities and legitimize the harms they cause.

A good example of such neutralizing appearance of procedural fetishism, which nonetheless reinforces global inequalities, is found in digital constitutionalism’s blind eye to very important substantive policy issues, such as climate change, environmental degradation, and continued exploitation of resources from so-called third world countries.

61. Steffek, supra note 56, at 784.
The EU- and US-dominated digital constitutionalism has focused on inequalities reinforced by tech companies in organizational settings in business and public administration. It has largely been blind to the inequalities on a global scale, including global outsourcing of labor and flow of capital through colonial and extractive processes. While industrial nations in North America, Europe, and East Asia compete in the tech sector, digital technologies depend on global resources and underpaid, racialized labor, most often extracted from the so-called Global South. Critical scholarship has demonstrated how the profits for big tech companies depend on large-scale exploitation of soil, minerals, and other resources. Others have argued that the process of extraction and exchange of personal data is itself a form of dispossession and data colonialism.

B. Legitimizing Effect of Procedural Fetishism

Moreover, procedural fetishism is used strategically not only to distract from power disparities but also to legitimize unjust and harmful substantive policies and actions by exploiting perceptions of legitimacy and justice. As early as the 1980s, psychological research undermined the traditional view that substantive outcomes drove perception of justice by showing that it was more about the procedure for reaching the substantive outcome. In the same way, procedural fetishism is used to strengthen the perceived legitimacy of tech companies without any

65. See Padmashree Gehl Sampath, Governing Artificial Intelligence in an Age of Inequality, 12 GLOBAL POLY 21 (2021).
70. Id. at 17.
meaningful change in their substantive policies or accountability mechanism. Procedural fetishism conveniently exploits the human tendency to accept unjust decisions by giving the (misleading) impression that people provide meaningful input. Many of the ongoing proceduralist reforms of the digital constitutionalist project, such as Facebook’s Oversight Board, are primarily conceived to make it look like Facebook is doing the “right thing” and delivering justice, regardless of whether substantive policy issues change or not. Importantly, such corporate initiatives divert attention from the power problems caused by the global dominance of the tech companies. The language of “lawfulness” and constitutional values, prevalent in digital constitutionalism, works as a particularly strong and legitimizing catalyst both in public and policy debates. As critical scholars have pointed out, using the terminology, which is typically employed in the context of elected democratic governments, misleads for it infuses tech companies with democratic legitimacy and conflates corporate interests with public objectives.

C. Procedural Fetishism and Administrative Law

Yet, procedural fetishism is not unique to the corporate sector or to a particular form of political organization (e.g., democracy). Rather, it is closely intertwined with the origins of the modern state. Indeed, administrative law, at the heart of modern public administration, displays similar excessive focus on procedural form in containing government power as compared to the substance of government policy and action. For example, in their recent book Law & Leviathan: Redeeming the Administrative State, American legal scholars Cass Sunstein and Adrian Vermuele proposed a set of procedural principles “with widespread appeal in many legal systems, . . . often discussed under the heading of natural justice, natural procedural justice, or . . . inherent in the notion of ‘due process of law.’” These principles would constitute the morality of administrative law: (1) agencies must follow their own rules; (2) retroactive rulemaking is disfavored and must be

limited to prevent abuse; and (3) official agency declarations of law and policy must be congruent with the rules that agencies actually apply.76

Nicolas Bagley has described the proceduralist obsession in administrative law as a “procedure fetish,” where the legitimacy of government practice is assessed without regard to its substance—“how well it advances [society’s] collective goals”—but solely with regard to the form the practice takes or the procedures governing it, “the stringency of the constraints under which it labors.”77 This fetishization of procedure, Bagley argues, is necessary to legitimize the administrative state to the skeptical public.78

In the following sections, I suggest that procedural fetishism plagues digital constitutionalism not accidentally but through sustained corporate and state efforts.

IV. PROCEDURAL FETISHISM AND STATUS QUO: CORPORATE ACTION & HELP FROM THE STATE

Tech companies vigorously promote procedural fetishism to legitimize their power and business models, to avoid regulation, and to enhance their reputation for commercial gain. Tech companies often come up with private solutions to develop apparent safeguards against their own abuse of power and to increase their transparency to the public. Yet, many such corporate initiatives are designed to obfuscate and misdirect policymakers, researchers, and the public in the bid to strengthen their brand and to avoid regulation and binding laws.79 In the following sections, I will show that tech companies have also successfully corporatized the laws and regulations that bind them. Through many procedures, checklists, and frameworks, corporate compliance with existing binding laws has often been a strategic performance, devoid of substantial change in business practices. Such compliance has worked to legitimize business policy and corporate power to the public, regulators, and the courts. In establishing global dominance, tech companies are also aided by governments, policymakers, and the courts.

A. Procedural Washing Through Self-Regulation

I will first look at self-regulation, voluntary commitments, and CSR,

76. SUNSTEIN & VERMEULE, supra note Error! bookmark not defined., at 9.
on which our digital constitution has often relied. Despite their popularity, voluntary corporate initiatives are impotent to realize the digital constitutionalist project and constrain the power of tech companies. These initiatives merely promote corporate interests of avoiding regulation and furthering profits.

1. Self-Regulation as Branding, Risk, and Management Tool

First, corporate self-regulatory initiatives are often cynical marketing and social branding strategies to increase public confidence in operations and create a better public image. CSR scholars have long argued that voluntary corporate initiatives have been captured by business interests and commodified. Tech companies self-regulate selectively by disclosing what is commercially desirable to disclose and address. For example, Google, when creating an Advanced Technology External Advisory Council in 2019 to help the company to implement Google’s AI Principles, simultaneously refused to reveal the internal processes that led to the selection of a controversial member, Kay Coles James, an anti-LGBTI advocate and climate change denial sponsor. While employees’ activism forced Google to rescind the council, ironically, this showed Google’s unwillingness to publicly share the

---


85. In this sense, this episode relates also to the necessity to stabilize self-contestation mechanisms within private and hybrid regimes and organizations, in order to effectively democratize them, as argued in societal constitutionalism scholarship. Cf. Gunther Teubner, Quod omnes tangit: Transnational Constitutions Without Democracy?, 45 J. OF L. AND SOC’Y S5 (2018).
selection criteria of their transparency boards. Internationally, a powerful platform for what Ronen Shamir called “corporate-inspired versions of social responsibility” has been created by the UN Guiding Principles. Now more than ten years in action, the principles have received a lot of criticism from both scholars and leading NGOs for simply “endors[ing] the status quo: a world where companies are encouraged, but not obliged, to respect human rights.” However, empirical research by the organization Rights and Accountability in Development (RAID) suggests that corporations endorse the UN Guiding Principles because they “offer companies a way to manage human rights risks, thereby protecting their business reputation, insuring against claims, and managing problems to avoid their escalation. Ultimately, like any other risk management process, it is an approach which protects profits by reducing costs.”

This is in line with proceduralist arguments that “mandating procedures is more politically viable than substantive interventions.” For example, tech companies have realized that dispute resolution procedures improve profits. Corporations adopt company-based grievance mechanisms to overcome barriers to accessing judicial review while at the same time introducing numerous controls to monopolize information, such as legal waivers and confidentiality clauses. This channels victims through a review mechanism of the company’s own making, which is centrally devised and controlled. Thus adopting and implementing the UN Guiding Principles-inspired human rights impact assessments, due diligence, and organization-based resolution mechanisms do not change much in substance and sometimes could make things worse, such as undermining or completely denying access to remedy. Thus, private proceduralist initiatives are often used

86. RONEN SHAMIR, supra note 81, at 105.
90. See Van Loo, supra note 6, at 894.
92. See RIGHTS AND ACCOUNTABILITY IN DEVELOPMENT (RAID), supra note 89, at iii–vi.
strategically and cynically to lower the risks and manage their brand without substantial changes in business policies and practices.

2. Voluntary Commitments vs. Profit

Second, corporations participate in CSR frameworks only if it pays off for them in the long run. Classical CSR theories suggest that consumer demands for corporate responsibility, “naming and shaming” practices by NGOs, pressure from socially responsible investors, and actual consumer boycotts may push private actors to abandon certain policies or values and adopt others. While profitability is not the only reason driving corporations to embrace voluntary commitments, it is nonetheless the most influential. Proponents of proceduralist interventions in platform governance agree that tech companies have adopted dispute resolution procedures because they realized that such procedures improve profits. When commitments to social values and profitability conflict, the latter will often prevail. This is well illustrated by the complicity of Google in restricting free speech in China. An enormous public outcry has not been enough to reverse agreements made by Google to return to China to expand its customer base. While Google’s commitment to human rights was questioned by many, even a special “China search database” does not prevent Google from branding itself as a defender of “internet freedom.” In this sense, proponents of proceduralist interventions are right that mandating procedures is more politically viable than substantive interventions. Yet, it is not a strength, as proceduralists see it, but rather a weakness of digital constitutionalism.

A similar corporate appetite for profits can be seen in IBM’s supply of facial recognition technology to law enforcement. For example, in 2012, IBM provided police forces in the Philippines with video


95. See Van Loo, supra note 91, at 555–56.


98. See Van Loo, supra note 6, at 894.
surveillance software, which was used to perpetuate President Duterte’s extrajudicial killings in his “war on drugs.” At the time, IBM defended the deal with the Philippines, saying it “was intended for legitimate public safety activities” but claimed that it had ceased the provision of its technology to the Philippines in 2012. Yet, IBM mentioned its “Face Capture” technology in a public disclosure in 2013 and 2014 presentations on its Davao City project. The company’s practice of providing authoritarian regimes with infrastructure dates back to the 1930s when IBM supplied the Nazi Party with unique punch card technology that was used to run the regime’s surveys to target Jewish people.

3. Self-Regulation for Preventing Regulation & Corporate Positioning as Regulators

Third, corporate initiatives also allow tech companies to prevent regulation of their activities. A good example is Facebook’s Oversight Board, established in May 2020. With forty members from “a variety of cultural and professional backgrounds,” the role of the board is what Mark Zuckerberg has described as a quasi-“Supreme Court” to apply Facebook’s policies and assess public interest considerations to overrule content moderation decisions. According to the Oversight Board’s website, “[t]he purpose of the board is to promote free expression by making principled, independent decisions regarding content on Facebook and Instagram and by issuing recommendations on the relevant Facebook company content policy.” Proponents claim that Oversight Board members have so far used their policy recommendations strategically to “talk” to a broader audience in attempts to change Meta’s corporate policy.

100. See id.
101. See id.
However, an establishment of an oversight board does not per se result in any substantial changes. It only reviews individual decisions rather than overarching policies. While some decisions of the Oversight Board, such as the one about Trump, have increased the scrutiny and social pressure on Meta, which also has, in some instances, followed up on decisions, attention remains not on critiquing the legitimacy or appropriateness of Facebook’s business practices themselves but rather on Facebook’s “procedures” and “safeguards” about them. The appropriateness of the underlying business model is obfuscated, or even legitimated, through the microprocedural initiatives, with little power to change the status quo.

In setting up the board, Facebook has attempted not only to stave off regulation but also to position itself as an industry regulator by inviting competitors to use the model of the Oversight Board as well. The weak and almost nonexistent regulation under our Digital Constitution has further solidified the move away from “territorial to functional sovereigns,” where companies position themselves as decision-making authorities that exercise juridical power. This move results in a shift of the identity of digital companies from participants in the market to makers of the market, creating a new mode of political economy and concentrating power in the hands of corporations rather than governments. Companies can then depict themselves as their own regulators, as Facebook has done through its quasi-judicial Oversight Board.

4. Self-Regulatory Initiatives to Reinforce the Status Quo & Inflict Further Damage

Many corporate initiatives and voluntary instruments are not

---


merely insufficient in tempering corporate power. They cement and legitimize the status quo where corporations can continue with their business models without changing the substance of their practices, often even framing themselves as regulators, as Facebook has done through its quasi-judicial Oversight Board. For example, the UN Guiding Principles simply cement the status quo that corporations are not obliged to follow their human rights commitments and that the corporate role in relation to fundamental rights is itself a very limited one that needs to be distinguished clearly from that of the state.\textsuperscript{112} In addition to reinforcing the status quo, voluntary initiatives and CSR principles often do further damage in regulating corporate behaviour. For example, the UN Guiding Principles have successfully replaced “obligation” —a term with clear legal meaning under international law—with the much fuzzier concept of “responsibility.”\textsuperscript{113} However, if we wish to constrain the power of tech companies, corporations should have binding obligations under both national and international law.\textsuperscript{114}

\textbf{B. Procedural Washing Through Law & Help of State}

Have the laws of digital constitution performed better than self-regulation? I argue that even as digital constitutionalist efforts via national or regional laws have become more prevalent in recent years, tech companies have exploited the ambiguity of laws regulating their behaviour through performative compliance, devoid of changes in their core practices and business models. Often, policymakers have compounded this problem by creating legal provisions to advance the proceduralist agenda of corporations, including via international organizations and international law. Regulators and courts have enabled corporatized compliance in applying these provisions by focusing on the quality of procedural safeguards.


\textsuperscript{113} Robert Blitt, \textit{Beyond Ruggie’s Guiding Principles on Business and Human Rights: Charting an Embracive Approach to Corporate Human Rights Compliance}, 48 TEX. INTL L. J. 33, 44 n. 65 (2012) (“The plain meaning of ‘responsibility’ suggests a moral obligation to behave correctly or a thing that one is required to do, rather than a duty to which an actor is legally bound. Although the final Guiding Principles do not provide explicit recognition that ‘responsibility’ is distinct from ‘duty’ or ‘obligation,’ the difference is implied insofar as the term duty is invoked in regard to states only.”).

1. Corporate Procedural Washing Exploits Ambiguity of Law

To illustrate this argument, I will look at data protection and privacy law—an area of human rights law that has been applied horizontally to private actors. The corporate exploitation and undermining of the General Data Protection Regulation (GDPR)\textsuperscript{115}—the piece of legislation that has gained the reputation as the strongest and most ambitious law of the digital constitution through symbolic compliance—is of particular interest. For example, the GDPR requires “unambiguous” consent for collection of ordinary, nonsensitive data,\textsuperscript{116} and “explicit” for sensitive data, health information, sexual orientation, and racial and ethnic background.\textsuperscript{117} However, neither the GDPR nor any EU guiding documents clarify how consent can be made explicit or unambiguous, or how these requirements differ in practice.\textsuperscript{118} Similarly, in the United States, the FTC requires corporations to provide “adequate notice to consumers” without providing specific guidelines how the “adequate” notice should be delivered.\textsuperscript{119} As Ari Ezra Waldman has demonstrated, these ambiguous legal requirements are interpreted by legal compliance professionals in accordance with corporate values rather than the substantive goals of the law.\textsuperscript{120}

Similarly, procedure-focused provisions in data privacy law establish certain safe harbours, which enable the companies to avoid liability by simply following a prescribed procedure. For example, under the traditional notice-and-consent regime in the United States, companies avoid liability as long as they post their data use practices in a privacy policy.\textsuperscript{121} Similarly, under the GDPR, the so-called data protection impact assessments can shield companies from liability even when engaging in “high risk” data processing activities that pose danger.

\textsuperscript{115} Regulation 2016/697, of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) [hereinafter GDPR].

\textsuperscript{116} Id. art. 4 at 11.

\textsuperscript{117} Id. art. 9, at 2(a).


\textsuperscript{120} Waldman, supra note 117, at 797.

\textsuperscript{121} Joel R. Reidenberg et al., Disagreeable Privacy Policies: Mismatches between Meaning and Users’ Understanding, 30 BERKELEY TECH. L.J. 1, 41 (2015).
to individuals.  

2. Corporatization and Managerialization of Law

The regulatory regime of data privacy is “managerialize[d]”: it is interpreted by compliance professionals, human resource experts, marketing officers, outside auditors, and in-house and firm lawyers, as well as systems engineers, technologists and salespeople to prioritize values of efficiency and innovation in the implementation of privacy law. For example, the third-party privacy technologists, who are not legally trained, are hired to design and implement an organisation’s privacy regime, bringing along their “managerial-compliance” focus and dictating how a corporation should comply. Chief Privacy Officers perceive themselves as “stewards of data and being responsible to consumers,” whilst the systems engineers make decisions about design integration and system efficiency. Often the individuals performing these roles do not collaborate at all. Many symbolic structures of compliance are created, yet, apart from an exhaustive suite of checklists, toolkits, privacy roles, and professional training, there are hardly substantial actions to enhance consumer protection or minimise online data breaches.

These structures comply with the law in name but not in spirit, which is treated in turn by lawmakers and judges as best practice. The law thus fails to achieve its intended goals as the compliance metric developed by corporations becomes dominant. Waldman has termed this process legal endogeneity, a phenomenon where law is based on the ideology of the actors whose actions it seeks to regulate, resulting in “mere symbols of compliance . . . standing in for real privacy protections.” Procedurally oriented approaches to regulating tech companies fail, because the “mere presence of compliance structures” is

122. GDPR, supra note 114, art. 35 at 1.
123. Waldman, supra note 117, at 778.
124. Id. at 45, 50–58 (discussing third-party technologists like Cybersaint and Nymity based in the United States).
128. Waldman, supra note 117, at 777 (citing LAUREN B. EDELMAN, WORKING LAW: COURTS, CORPORATIONS, AND SYMBOLIC CIVIL RIGHTS (2016)).
129. Waldman, supra note 117, at 776.
130. Id.
assumed to be “evidence of substantive adherence with the law.”\textsuperscript{131} This is similar to the model proposed by the UN Guiding Principles, discussed above.

Twenty-six recent studies analyzed the impact of the GDPR and US data privacy laws and none have found any meaningful influence of these laws on data privacy protection of the people.\textsuperscript{132} While data privacy is just one policy area of the digital constitution—similar exploitation of the ambiguity and managerialization in the law must occur in other areas. Indeed, the law itself has often been designed in the spirit of procedural fetishism, enabling corporations to avoid liability and change their substantive policies by simply establishing proscribed procedures.

3. Procedural Fetishism of Regulators and Courts

Regulators and the courts, by emphasizing procedural safeguards, also engage in performative regulation, grounded in procedural fetishism, that limits pressure for stricter laws by convincing citizens and institutions that their interests are sufficiently protected without inquiring into the substantive legality of corporate practices.\textsuperscript{133} A good example of proceduralism in regulation is FTC audits and “assessment” requirements, which require corporations to demonstrate compliance through checklists.\textsuperscript{134} Similar procedural fetishism is also prevalent in jurisprudence. For example, the Court of Justice of European Union (CJEU) exerts a prominent role in regulating tech companies. Its recent judgments focus on the platforms’ responsibility for removing harmful content,\textsuperscript{135} the passivity and specificity of consent required for cookies,\textsuperscript{136} and the geographical scope of the (in)famous “right to be

\textsuperscript{131} Id. at 792–94.
\textsuperscript{132} Filippo Lancieri, Narrowing Data Protection’s Enforcement Gap, 74 ME. L. REV. 15, 16 (2022).
\textsuperscript{133} Ari Ezra Waldman, Privacy, Practice, and Performance, 110 CAL. L. REV. 1221, 1234 (2022).
\textsuperscript{134} For an excellent account of FTC work on privacy, see Chris Jay Hoofnagle, Federal Trade Commission Privacy Law and Policy 167 (2016).
forgotten.” And it seems that the CJEU is delivering judgments of substantive limits on tech companies. However, closer analysis of the judgments reveals that often they change little in corporate practices. For example, the CJEU has recently developed the “staged” allocation of responsibility in joint controllership of personal data, in its own words, aiming to leave “no gaps” in the protection of individuals. However, “fragmenting” controllership into different stages of data processing again obscures the bigger picture by emphasizing microissues. Such fragmentation is incompatible with the GDPR, for it does not reveal the intended purposes of data processing, and thus negates informed and specific consent. Such fragmentation obscures the pervasiveness of data commodification in the digital economy. This is similar to what I call the procedural fetishism of European mass surveillance jurisprudence. This jurisprudence does not assess state surveillance practices by reference to their effectiveness in advancing the proclaimed goals of enhancing national security, but rather purely to the stringency of the procedures governing the surveillance programmes. The political pressure on both the CJEU and the European Court of Human Rights (EctHR) from member states, EU organs and major external actors like the United States has led to the courts’ emphasis on “robust procedures” rather than the substantive legitimacy (the “ends and standards” the regimes pursue).

4. Procedural Washing Through State Rhetoric and International Law

Procedural washing by corporations has also been aided by executive governments in their efforts to establish and sustain global dominance—both through large amounts of public funding and subsidization to these companies and through the development of the laws, including international laws, that suit the corporate and national agenda. Such support is not one-sided, of course, as the state expands


140. Zalnieriute, supra note 64, at 185–92.
its surveillance and geopolitical power through technology companies. All major powers, including the United States, the European Union and China, have been active in promoting their tech companies. A mutually beneficial and interdependent relationship between the US government and US tech companies has been studied by Herbert Schiller, who described the continuous subsidization by US companies of private communications companies back in the 1960s and 1970s. Building on these insights, Powers and Jablonski describe how the dynamics of what they call the information-industrial-complex have catalyzed the rapid growth of information and communication technologies within the global economy while firmly embedding US strategic interests and companies at the heart of the current neoliberal regime. Such central strategic position necessitates continuous action and support from the US government.

To maintain the dominance of US tech companies internationally, the US government opposes and has so far effectively prevented any international organization or multilateral treaty over data privacy. It also aggressively promotes the global free trade regime, intellectual property enforcement, and other policies that suit US interests. For example, the dominance of US cultural and information technology products and services worldwide is secured via the free flow of information doctrine at the World Trade Organization, which the US State Department pushed with the GATT, GATS and TRIPS agreements. The free flow of information doctrine is a crucial element in the US version of digital constitution because it allows US corporations to collect and monetize personal data of individuals from around the world. This way, data protection and privacy are not part of the “universal” values of the internet, whereas strong intellectual property protection, or a ban on gambling, is not only viable and doable but also strictly enforced globally. Against the background of continuous political support on an international level and subsidization of tech companies on a domestic level, the weakness of digital constitutionalism (e.g., direct access by US government and agencies

144. Herbert I. Schiller, Culture, Inc.: The Corporate Takeover of Public Expression 118 (1991); Schiller, supra note 141, at 93.
145. This is beyond the scope of this article, but see Marianna Mazzucato, The Entrepreneurial State: Debunking Public vs. Private Sector Myths (1st ed. 2013).
such as NSA to the tech companies’ infrastructure), as revealed by Edward Snowden back in 2013, is not surprising.

While the role of US government in shaping the weak digital constitution has been dominant, many other governments have been complicit. The EU, which is often touted as the leader in the regulation of digital technologies, is dependent on the US government for surveillance infrastructure. It is true that European Digital Constitutionalism has attempted, at least on the surface, to constrain US tech giants through European regulatory and enforcement efforts. However, such efforts are underlined by motives to strengthen its own tech industry and dominate global markets.footnote{146} For example, the EU Artificial Intelligence (AI) Act, despite its declared mission to “human centered AI,” is silent about the environmental degradation and social harms that occur in other parts of the world because of large scale mineral and resource extraction and energy consumption, necessary to produce and power AI and digital technologies.footnote{147} The EU AI Act is also silent on the conditions under which AI is produced and the coloniality of the AI political economy. It does not address precarious working conditions and global labor flows.

Thus, at least in the United States and the European Union, digital constitution has favored the status quo. It has prioritized inaction, that is, self-regulation over regulation; no enforcement over enforcement; and judicial acceptance over substantial resistance. While regulatory strategies in the United States and the European Union differ and are often framed as in stark opposition or in competition to set the global standards (e.g., the EU has stringent comprehensive data privacy regulation for private actors, whereas the United States has not), the end result has been similar. Stressing the differences in US and EU regulatory approaches has been popular. It enables us to speak about distinct cultural traditions of privacy,footnote{148} different neoliberal or ordoliberal conceptions of economic constitutionalism and market


footnote{147} Mark Coeckelbergh, AI for Climate: Freedom, Justice, and Other Ethical and Political Challenges, 1 AI ETHICS 63, 70 (2021); Payal Dhar, The Carbon Impact of Artificial Intelligence, 2 NATURE MACH. INTEL. 423, 423–25 (2020); Emma Strubell et al., Energy and Policy Considerations for Modern Deep Learning Research, 34 PROC. AAAI CONF. A.I. 13693 (2020).

against Procedural Fetishism

It also gives us a feeling that regulation of tech companies is fundamentally different in the United States and the European Union. However, the end result has been very similar both in the European Union and the United States: the tech companies collect and exploit personal data not only for profit but also for political and social power.

In sum, digital constitution, plagued by procedural fetishism and colonial and imperial ambitions, is dangerous for creating an illusion that it is normatively neutral. However, it merely obscures the normative values that underly corporate, regulatory, judicial actions under its veil of neutrality. Corporate strategies to avoid regulation or performative, symbolic compliance with the laws enable the legitimation of corporate power and rule on a macro level through its emphasis on procedure—a sort of micro critique. And our normative values of constraining tech companies are replaced with the corporate vision of division of power and wealth between the corporations and the people, masked under neutrality.

V. A New Digital Constitution: Rectifying Power Imbalance

To rectify the imbalance of power between people and tech companies, a new digital constitution thus must try something different: it must shift its focus from soft law initiatives and performative corporate compliance to substantive accountability and tangible legal obligations by the tech companies. Imposing directly binding data privacy obligations on tech companies with an international treaty is one (but not the only) option to take the digital constitutionalist project further. Other parallel actions include breaking up and taxing tech companies, increasing competition and regulatory enforcement, increasing public scrutiny, and democratizing big tech, in other words, involving people in the governance of tech companies. With all these mutually reinforcing efforts, the digital constitution must debunk procedural fetishism and demand basic tangible constraints for the new social contract between people and tech companies. Under the new digital constitution, the legitimacy of corporate action, as well as our law trying to tackle their power, must be assessed without regard to procedures corporations follow or new initiatives they establish.

Instead, our laws and corporate policies must be assessed on how well they achieve public goals; how fairly they distribute power and wealth; how well they protect from corporate predations; and how well they minimize the risks to people, health, and welfare.

A. Increased Political Scrutiny & Novel Ambitious Laws

The first element of the new digital constitution must be an increased political scrutiny of tech companies and the legislative regimes that facilitate their growth. Tech companies came under increased scrutiny after the Snowden revelations in 2013, a process that was reenergized by the Cambridge Analytica scandal, the US Election interference scandals in 2016, and the Covid-19 pandemic in 2020. The same scrutiny should be applied to legislative regimes that have in the past established safe harbours for tech companies, and new laws should be implemented to eliminate procedures that legitimise harmful business practices. For example, legal frameworks, such as EU-US Transatlantic Privacy Framework,\(^{150}\) previously known as Safe Harbour and Privacy Shield, should be abolished. Similarly, regimes, based on procedural avoidance of liability, such as the one under Section 230 of the US Communications Decency Act 1996, should be reconsidered. This provision provides that websites are not treated as the publisher of third-party content (i.e., user submitted content) and is particularly useful for platforms like Facebook. Former Speaker of the House of Representatives Nancy Pelosi recently described it as a “gift” to tech companies and did not “think that they are treating it with the respect that they should.”\(^ {151}\) Federal bills like Health Misinformation Act\(^{152}\) or Platform Accountability and Consumer Transparency (PACT) Act \(^{153}\) suggest Section 230 is under pressure. US Democrats claim platforms are not acting responsibly with regard to health misinformation, while


Some of the more recent laws of the digital constitution might be displaying the first seeds of substance-focused regulation. For example, more recent data privacy legislation, such as the California Consumer Privacy Act of 2018, include prohibitive requirements, such as the requirement for a “Do Not Sell My Personal Information” link.\footnote{155. California Consumer Privacy Act of 2018, CAL. CIV. CODE § 1798.135(a)(1) (2018).}{155}

Several other US states, including Colorado, Virginia, and Utah, have also recently enacted comprehensive data privacy laws, which will come into effect in 2023.\footnote{156. Colorado Privacy Act, COLO. REV. STAT. S.B. 21-190 (2021) (effective July 1, 2023); Virginia Consumer Data Protection Act, VA. CODE ANN. § 59.1-575 (2021) (effective Jan. 1, 2023); Utah Consumer Privacy Act, UTAH CODE ANN. S.B. 227 (2022) (effective Dec. 31, 2023); see also IAPP Resource Centre, supra note 14.}{156}

Another example is the many moratoriums issued on the use of facial recognition technologies across many municipalities and cities in the United States, including the state of Oregon and New York City.\footnote{157. Monika Zalnieriute, \textit{Burning Bridges: The Automated Facial Recognition Technology and Public Space Surveillance in the Modern State}, 22 COLUMBIA SCIENCE AND TECH. REV. 314 (2021).}{157}

While these bans have focused on law enforcement, there is no reason why similar logic could not be extended to private power too.

In the EU too, some of the latest proposals also display an ambition to ban certain uses and abuses of technology. For example, unlike the GDPR or the Digital Services Act, the Artificial Intelligence Act provides a list of “unacceptable” AI systems and prohibits their use. The Artificial Intelligence Act has been subject to criticism about its effectiveness,\footnote{158. See, e.g., Michael Veale & Frederik Zuiderveen Borgesius, \textit{Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach}, 22 COMPUT. L. REV. INT'L 97 (2021); Vera Lúcia Raposo, \textit{Ex Machina: Preliminary Critical Assessment of the European Draft Act on Artificial Intelligence}, 30 INT'L J. L. AND INFO. TECH. 88 (2022); Lilian Edwards, \textit{Expert Opinion: Regulating AI in Europe. Four Problems and Four Solutions} (Mar. 31, 2022), https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/.}{158}

yet its prohibitive approach can be contrasted with earlier EU regulations, such as GDPR, which did not exclude any areas from data processing. On an international level, the OECD has recently announced a landmark international tax deal, where 136 countries and jurisdictions representing more than 90 percent of global GDP agreed to a minimum corporate tax rate of 15 percent on the biggest international
corporations which will be effective in 2023.\textsuperscript{159} While this is not tackling tech companies’ business practices, it is aimed at fairer redistribution of wealth, which too must be the focus of the new digital constitution.

Recent attempts to ban behavioural advertising in the Digital Services Act have failed,\textsuperscript{160} but the laws of the new digital constitution should aim to limit the power of online platforms to collect, combine, and exploit data for the purposes of delivering targeted advertising as well as for developing and improving platforms’ services. As argued by civil society advocates, the Digital Services Act (DSA) package should introduce specific obligations and prohibitions for online platforms, that is, impose a clear standard of interoperability and prohibit practices and behaviours that impede access to the market for competitors.\textsuperscript{161} In short, the novel and ambitious laws of our new digital constitution should promote alternative funding models, such as contextual advertising and subscription and public subsidies, rather than accepting corporate business models as inevitable.

\textbf{B. Regulatory and Punitive Enforcement}

Another crucial component in furthering the ambition of the digital constitution is regulatory and punitive enforcement action against tech companies. For example, Google’s market dominance saga and Facebook’s Cambridge Analytica scandal suggest that private actors will rarely change their policies and procedures unless threatened with direct legal and punitive actions by influential institutions, such as the EU Commission or the US Department of Commerce. Many commentators claim that underenforcement, especially by regulators in Ireland, has undermined the effectiveness of the GDPR.\textsuperscript{162} A recent leak


\textsuperscript{160} KAROLINA IWAŃSKA, To Track or Not to Track? Towards Privacy-Friendly and Sustainable Online Advertising, PANOPTYKON FOUNDATION (Nov. 2020), https://panoptykon.org/sites/default/files/publikacje/panoptikon_to_track_or_not_to_track_final.pdf.


suggests that Facebook has been operating in violation of the GDPR. Yet, recent enforcement actions brought by the United States and the European Union show that tech companies are under intensifying regulatory pressure. For example, in 2020, the US House Antitrust Subcommittee issued a damning report calling out Facebook for anticompetitive practices, and the FTC is continuing its investigations into the company’s practices. Enforcement should also tackle unequal distribution of power, not just specific corporate conduct, and it seems that some of the new laws are heading that direction. For example, the Digital Service Act gives extensive oversight powers to the commission and national Digital Service Coordinators, such as requesting disclosures and conducting on-site inspections. Regulatory enforcement alone is not sufficient to shift from procedural fetishism to substantive regulation and limits on private power but is a necessary component of the digital constitution, supplementing other tools.

C. Breaking Up Tech Companies

The new digital constitution must also break up tech companies, many of which have grown so large that they are effectively gate keepers in their markets. Many scholars have recently proposed ways to employ antitrust and competition law to deal with or break up big tech companies, and such efforts are also visible on a political level. For example, in December 2020, the EU Commission published a proposal for two new pieces of legislation: the Digital Markets Act (DMA) and the

---


Digital Services Act (DSA). The proposal aims to ensure platform giants, such as Google, Amazon, Apple, and Facebook operate fairly and to increase competition in digital markets. A central feature in the DMA is the concept of “gatekeepers,” which is defined as a provider of core platform services that: (i) has a significant impact on the internal market (i.e., has Union turnover bigger than EUR 7.5 billion or a market cap equal or bigger than EUR 75 billion); (ii) acts as an important gateway for businesses to reach end users; and (iii) enjoys an entrenched and durable position in its operations. Several obligations will apply to gatekeepers, including allowing business users to offer the same product or service to end users through third-party intermediation services and allowing business users to promote offers to end users acquired via the core platform service. The DMA also imposes prohibitions on gatekeepers, including treating services offered by the gatekeeper more favourably than those offered by third parties, preventing consumers from linking up to business outside of gatekeeper platforms, and preventing users from uninstalling preinstalled software.

The antitrust movement—known as neo-Brandeisianism in the United States—and the competition law attempt in the EU have been part of the digital constitutionalist project for a while, yet they have been of limited success so far. Legal instruments for breaking up the concentration of power in the tech sector exist. For example, the US Sherman Antitrust Act of 1890 makes monopolization unlawful. However, the will to regulate and break up tech companies in the United States, where their dominance has been useful for the US government and economy, has been limited. For example, in the late 1990s the US Department of Justice targeted Microsoft for tying its web browser to its operating system, and Microsoft was ordered to break it up into two separate companies, one for operating systems and another.


170. Id. arts. 5–6.

171. Id.

for application software.\textsuperscript{173} However, this decision was reversed and Microsoft was given behavioural restrictions, including the licensing of communication protocols and application programming interfaces as well as prohibitions on blocking “middleware.”\textsuperscript{174}

In contrast to this weak political will, our new digital constitution must use the tools of competition and antitrust law (but not only them) to redistribute wealth and power. For instance, as Kieron O’Hara suggested, we could prevent anticompetitive mergers and require tech giants to divest companies they acquired to stifle competition, such as Facebook’s acquisition of WhatsApp and Instagram.\textsuperscript{175} We could also ring-fence giants into particular sectors. For example, Amazon’s purchase of Whole Foods Market (a supermarket chain) would likely be prevented by that strategy. Finally, we can force tech giants to split their businesses into separate corporations. As Zephyr Teachout suggests, instead of boycotting tech giants, we should demand that antitrust authorities split them up.\textsuperscript{176} For instance, Amazon would be split into its e-commerce platform, physical stores, web services, and advertising business.

\textbf{D. Beyond Breaking: The Public Utilities Approach}

However, other measures are needed to make the new digital constitution successful. While antitrust and competition proposals can be the starting point to wider criticism of tech companies, their dominance has marginalized other solutions for taming the power of tech companies in both public discussion and digital constitutionalist discourse. An almost exclusive focus on competition reforms obscures more radical solutions which threaten the basic premises of markets, commodification, and corporate ownership. For example, as critical scholars have suggested, digital services could be conceived as public utilities: either as closely regulated private companies or as government run organizations, administered at municipal, state, national, or regional levels.\textsuperscript{177} Public utility proposals differ from the neo-

\begin{flushleft}


175. KIERON O’HARA & WENDY HALL, FOUR INTERNETS 119 (2021).

176. TEACHOUT, supra note 166.

\end{flushleft}
Brandeisian “Break ‘Em Up” logic, because public utility reforms aim at placing big tech companies under public control.

Public control solutions have been proposed with regards to telecommunications networks in the United States—for example, Bernie Sanders’s proposal for low-cost and subsidized public access to broadband internet was based on a public utilities approach. Other proposals in this area include a public cloud computing platform or even social media funded by government but operated by independent statutory organizations. For example, Michael Kwet notes a strong blueprint for a public utilities approach with social media in the Fediverse, a network of community-run Free and Open Source Software (FOSS)-based social networks, which interoperate without a single centralised owner. Interoperability across similar services is a crucial component of the public utilities approach, however, exact solutions will vary. For instance, platforms for ride sharing could be contained within geographic boundaries. A utilities approach provides a strong alternative to market-driven solutions to restore competition in the technology sector and has more potential to address the structural problems of exploitation, manipulation, and surveillance.

The power, influence, and size of tech companies and the markets they operate have all been constituted and enabled by the state, which provides the legal foundations for the tech companies’ activities. Our new digital constitution requires active state intervention only in regulating particular corporate decision-making procedures, but we need active state intervention in reshaping the legal foundations structuring the ownership, control, and governance of digital infrastructure, which enables digital markets.

E. Democratizing & Decolonizing Technology Infrastructure

Finally, the new digital constitution must “democratize” tech companies by giving users meaningful power to influence how company policies and practices should be developed. It is beyond the scope of this paper to articulate exact proposals here, but Hans Gersbach, for instance, has proposed voting procedures and a representative user council. Voting procedures can range from a simple one-vote-per-user

179. Powers, supra note 141, at 76.
system to a co-voting system, in which user votes are weighted and considered alongside shareholder votes. The second solution is a user council, which could consist of users with the highest usage frequency (“super-users”), the most expertise (“expert-users”), and ordinary users.

Importantly, such democratization of big tech must incorporate ideas from the so-called Global South and Indigenous epistemologies—they provide strong conceptual alternatives to dominant theories of technology and platform governance, which do have the potential to radically change the current regulatory landscape. Decolonial ways to theorize, analyze, and critique tech companies must be part of our new digital constitution because people in the so-called Global South relate very differently to major tech companies than those who live in the Global North. The “Technologies for Liberation” project studies how communities of queer, trans, two-spirit, black, and indigenous people are disproportionately impacted by surveillance technologies and criminalization and is a good starting point.

Creating and developing the new digital constitution requires novel spaces that are not dominated by procedural fetishism of our Western, often Anglo-Saxon position. A good example of possible resistance is a Non-Aligned Technologies Movement (NATM)—a worldwide alliance of civil society organizations, which aims to create “technosocial spaces beyond the profit-motivated model of Silicon Valley and the control-motivated model of the Chinese Communist Party.” NATM is modelled on the idea of the original collective of Non-Aligned Movement—a forum of 120 countries that are not formally aligned with or against any major power bloc. It sought to define an alternative to

---

182. Id. at 356.
capitalism and communism during the Cold War.

The democratization and decolonization must go beyond civil society or mere declarations from policy-makers. For example, the EU’s focus on regulating AI and ADMS as a consumer “product-in-use” in its AI Act, requiring individual protection, is hypocritical and undermines the claims to regulate “ethical” AI, for it completely ignores the exploitative practices and global implications of AI production and use. The new laws must recognize and explicitly acknowledge the power disparities, exploitation, and coloniality of data production as a collective rights deprivation and a new form of ongoing structural colonial violence.\(^{187}\)

In summary, digital constitutionalism should pay attention to the asymmetries in economic and political power on global scale. This covers both the US dominance in tech sector and US influence in shaping international free trade and intellectual property regimes. It also covers the rising influence of China and the EU’s ambitions to set global regulatory standards in many policy areas. Both business and public bodies in the so-called Global South are on the receiving end of Brussels demands of what “ethical” AI is, and how “data protection” must be understood and implemented.\(^{188}\)

VI. Conclusion

The new digital constitution must incorporate many different dimensions and strategies. They might not achieve substantive policy change alone. However, together and in parallel, they will enable us to start resisting corporate and state agendas of procedural fetishism. In the digital environment dominated by tech companies, procedural fetishism is an intentional strategy to obfuscate the implications of concentrated corporate power. Tech behemoths legitimize their practices through procedural washing and performative compliance to divert the focus onto the procedures they follow, both for commercial gain and to avoid their operations being tempered by regulation. They are also helped and assisted by states, which enable corporate dominance via the laws and legal frameworks.

Countering corporate procedural fetishism requires, first of all, returning the focus back to the substantive problems in the digital environment. In other words, it requires paying attention not only to the procedures but also to the substance of tech companies’ policies, practices, and power. This requires a new digital constitution. Rather

---

than buying into procedural washing as companies intend for us to do, we need a new digital constitution with binding, legally enforceable mechanisms to hold the tech companies accountable. We have many options, and we need to act on all fronts: we must redistribute wealth and power by breaking up and taxing tech companies, fortifying regulatory enforcement, increasing public scrutiny by adopting prohibitive laws, and democratizing and decolonizing big tech by giving people power to determine the way in which these companies should be governed. Our new digital constitution must recognize that tech companies exercise global dominance with significant international and environmental implications. This aspect of technology is related to global economic structure, and, therefore, cannot be solved alone: it requires systemic changes to our economy. The crucial step to such direction is developing and maintaining tech platforms as public utilities, which operate for the public good rather than profit. The new digital constitution should decommodify data relations, rethink behavioural advertising as the foundation of the internet, and reshape social media and internet search as public utilities. With all these mutually reinforcing efforts, the new digital constitution will debunk the corporate and state agenda of procedural fetishism and establish the new social contract for the digital age.
Tackling Threats to Academic Freedom Beyond the State: The Potential of Societal Constitutionalism in Protecting the Autonomy of Science in the Digital Era

RAFFAELA KUNZ

ABSTRACT:

While in the early days of the internet, there was great enthusiasm about the wealth of new possibilities for science, in recent years, awareness about the darker sides of digital technologies has been rising. It is known today that the leading academic publishers have tapped into the data analytics business. The resulting massive collection of user traces not only threatens privacy rights but also increases concerns about the consolidation of the oligopoly in the global academic publishing industry and large-scale corporate influence on science. This paper explores the consequences of this development for the constitutional protection of science. It argues that while classic accounts of constitutionalism are not entirely blind to constitutional challenges transcending the state-individual relationship, they struggle to capture the subtle yet systemic risks that the science system faces in the digital age. Societal constitutionalism is not only a useful lens to better understand these threats, but also to respond to them, providing valuable lessons for debates about digital constitutionalism and the effective protection of fundamental rights in the digital age.

Keywords: Academic Freedom, Open Access, Surveillance, Digital Constitutionalism, Fundamental Rights Theory.

* Postdoctoral Fellow, University of Zurich, Switzerland. raffaelamilena.kunz@uzh.ch.
I. INTRODUCTION

The internet offers infinite possibilities, but also new threats and pitfalls for science and academic freedom. On the one hand, the World Wide Web, for the first time, has made the global dissemination and nearly universal accessibility of knowledge possible. “An old tradition and a new technology have converged to make possible an unprecedented public good,” reads the enthusiastic opening of the “Budapest Open Access Initiative.” This document marks the beginning of the modern Open Access movement that emerged soon after the advent of the internet, demanding that academic publications are made freely accessible over the internet. On the other hand, in recent years, the awareness about darker sides of digital technologies is rising. Not only were commercial publishers successful in “harness[ing] the digital revolution into a counterrevolution,”—first by inventing the paywall and then by undertaking an economic re-interpretation of Open Access—but reports have furthermore begun to reveal that leading academic publishers have long tapped into the data analytics business. In other words, they began using the tools developed by big tech companies such as Google and Facebook to track the online behavior of researchers and collect massive amounts of data, making clear that science is “. . . no special case on the internet.”

This development is another case in point of how digital technologies give way to new forms of value extraction and can fundamentally alter the nature of goods, turning public goods into common-pool resources that need to be “managed, monitored, and protected, to ensure sustainability and preservation.” The claim that “knowledge, which can seem so ubiquitous in digital form, is, in reality more vulnerable than ever before,” today is truer than ever. The danger

in relation to the practice of “surveillance publishing”\textsuperscript{7} thus not only concerns privacy rights but is also a further step towards the enclosure and commercialization of science. In recent years, other sectors of the digital economy have shown how business models based on “surveillance capitalism,”\textsuperscript{8} in combination with other typical characteristics of digital platforms, may lead to “explosive growth”\textsuperscript{9} and “extreme scale economies.”\textsuperscript{10} The fact that similar developments now take place in science thus increases concerns about an emerging “knowledge industry”,\textsuperscript{11} the oligopolistic position of a handful of powerful publishers,\textsuperscript{12} large-scale corporate influence and the “platformisation”\textsuperscript{13} of science. “We are at a critical juncture,” reads the conclusion of a much-cited report on data tracking and the “growing trend of commercial acquisition of critical infrastructure in our institutions.”\textsuperscript{14} Additionally, the new version of the Budapest Declaration on Open Access, released at the occasion of the twentieth anniversary of the original document in early 2022, recognizes “certain systemic problems,”


\textsuperscript{9} OECD, \textit{AN INTRODUCTION TO ONLINE PLATFORMS AND THEIR ROLE IN THE DIGITAL TRANSFORMATION} 1 (2019).


\textsuperscript{13} Philip Mirowski, \textit{The Future(s) of Open Science}, 48 SOC. STUD. SCI. 171, 171 (2018).

listing “harms caused by proprietary infrastructure, commercial control of research access”, and “commercial control of research assessment indicators”, among other problems.\footnote{15}

While the consequences for free speech and the public sphere due to the increasing control of media and information infrastructure by a handful of powerful business actors are being discussed extensively today,\footnote{16} the consequences for the science system remain largely unexplored. The aim of this paper is to contribute towards filling this gap. It argues that the developments in science in the digital age risk to undermine the autonomy of the science system and thereby violate the fundamental right of academic freedom. While data tracking quite obviously causes tensions with the fundamental rights of individual researchers, this paper argues that the developments closely connected to digitality also threaten science as a \textit{system} or \textit{institution}. Methodologically, the paper explores the potential of the theory of societal constitutionalism—at the heart of which lies the idea to find answers to constitutional threats that transcend the nation state—in tackling these challenges. David Sciulli, whose work aims to address authoritarian drifts in modern societies, is considered to be the founding father of modern societal constitutionalism,\footnote{17} which in recent years has been significantly developed and refined—especially by Gunther Teubner—as a theory grounded in sociology, but linked to legal-normative perspectives.\footnote{18} While classic accounts of constitutionalism are not entirely blind to constitutional challenges transcending the state-individual relationship, they struggle to capture subtle, yet systemic, risks such as those the science system faces in the digital age. One of the core ideas of societal constitutionalism, strongly inspired by systems

\footnote{15. The \textit{Budapest Open Access Initiative: 20th Anniversary Recommendations, Budapest Open Access Initiative} (March 15, 2022), \url{https://www.budapestopenaccessinitiative.org/boai20/}.}


\footnote{17. See generally \textsc{David Sciulli, Theory of Societal Constitutionalism: Foundations of a Non-Marxist Critical Theory} (1992).}

TACKLING THREATS TO ACADEMIC FREEDOM BEYOND THE STATE 269

theoretical approaches,19 is that the inner logic of a given societal sub-

system needs to be taken into account to find effective answers to

threats. The theory thus not only promises to provide a useful lens to

grasp, understand, and describe constitutional threats in the digital age

but also to address them.

The paper proceeds in three steps. In the first step, it gives a concise

overview of the current developments in science, highlighting the dual

capacity of digital technologies to increase and restrict access to

information. While the existing discussions often focus on problematic

practices of the leading academic publishers, such as Elsevier, this

paper argues that this view is too short-sighted. A systems-theoretically

informed analysis shows that the dynamics in science are best

understood as interactions between different societal systems, each with

expansionist tendencies. Understanding these dynamics provides the

starting point for effective regulatory responses. In the second step, the

paper shows that these threats are constitutional threats. It discusses

the limits of existing constitutional doctrine in protecting the science

system from the threats emanating in the digital age, and the potential

of the theory of societal constitutionalism in filling protective gaps.

Finally, in the last part the paper turns to the question of how these

challenges can and should be addressed by public actors. To this end,

the paper will make some concrete policy proposals based on the theory

of societal constitutionalism. While the first part of the paper thus uses

societal constitutionalism as an analytical tool, parts two and three

engage with it as a normative theory.20

II. THE THREATENED AUTONOMY OF SCIENCE IN THE DIGITAL AGE

Before turning to the constitutional protection of science, we must

first consider the challenges and threats the science system today faces

in relation to digitality. Societal constitutionalism provides a useful lens

for this mainly descriptive exercise. This analysis shows that the

current developments have both external and internal dimensions. Most

obviously, the current developments evidence an increasing

commodification of science and an intrusion of actors of the digital

economy into the science system. Yet, a closer look reveals that the

dynamics can only be understood in combination with the self-

endangerment and the corruption potential of academics. Finally, to get

the full picture, a third actor needs to be considered: science policy


State Constitutionalism, 4 TRANSNAT'L LEGAL THEORY 44, 50–51 (2013).

20. On the distinction between the analytical and normative limbs of the theory, see

Golia & Teubner, supra note 18.
makers who increasingly get involved and play an active role in steering and shaping the developments in the science system. In other words, the successes of commercial publishers in science can only be explained because they resonate with other dynamics. Digital tools are not at the origin of these developments, but magnify, exacerbate, and accelerate preexisting problematic dynamics. Any answer to these challenges needs to account for this intertwine ment.

A. Academic Publishers: Towards a Platformization of Science in the Digital Era

Discussions about problematic trends in the contemporary science system usually mainly focus on academic publishers. Time and again, these powerful actors have been criticized, ranging from concerns about their oligopolistic position, their giant and yearly growing revenues—27 billion dollars in 2018, 28 billion dollars in 2019, and 26.5 billion dollars in the pandemic year 2020—and the profits that are significantly higher than in other sectors. Already the coming of age of the modern Open Access movement in the early days of the internet must be seen against the backdrop of the practices of these actors. The high prices for subscriptions to their journals, which had given rise to the “serials crisis,” in which even wealthy institutions in the “Global North” could not afford them anymore, were increasingly considered excessive and out of sync with the possibilities the internet offered.

The emergence of the modern Open Access movement can thus be read as a countermovement against the big academic publishers and the increasing “intellectual land-grab” and “enclosure” of science, with the aim to limit the excesses of the economic sphere, increasingly perceived as a threat to the autonomy and integrity of science. The idea to make research openly available via the internet seemed to better align with scientific processes than the predominant commercial model;

21. See Larivière et al., supra note 12.
Open Access was perceived to resonate with the fundamental scientific values such as the norm of *communism*, the first of the five Mertonian norms, according to which the results of research should not be private property, but instead be recognized as the fruit of collective scientific work. Because of its aim to push back external influence and preserve the autonomy of science, the emergence of the Open Access movement has even been considered a “constitutional moment” for the science system.

Today, it has become clear that the “access revolution” was not successful. Even though the transformation towards Open Access, further boosted by the COVID-19 pandemic in which “science has opened up in a unique way” while states closed their borders, is as far advanced as never before, it has not fundamentally altered the structures in place. While in the early days of the new millennium, many were convinced that the internet irrefutably would mean the end of traditional academic publishing, it has become clear that academic publishers were successful in adapting to the realities of the digital economy, firstly by inventing the “pay wall” and then by transforming Open Access into a business model. While there are different Open Access models, the most successful model is the so-called gold model which shifts the costs from readers to authors via “Article Processing Charges” (“pay to publish” instead of “pay to read”). The disadvantages of this shift are obvious—rather than excluding readers, this new


This is reflected, for example, in the 2003 Berlin Declaration on Open Access which precisely because of the potentially far-reaching effects on traditional publishing perceives Open Access not only as a chance, but also a challenge: “Obviously, these developments will be able to significantly modify the nature of scientific publishing as well as the existing system of quality assurance.” Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, MAX-PLANK-GESELLSCHAFT: OPEN ACCESS (Oct. 22, 2003), https://openaccess.mpg.de/Berlin-Declaration.
33. See WALT CRAWFORD, GOLD OPEN ACCESS 2016-2021: ARTICLES IN JOURNALS (GOA7) 1–2 (2022) (containing the latest numbers).
publication mode risks to exclude authors—all the more since numerous studies show that Open Access has not solved the price problem in the publishing system. Rather, it is likely that we are witnessing “an open access sequel to the serials crisis.”\textsuperscript{34} The effects are particularly detrimental for scholars in the “Global South” which risk to be excluded from participating in the scholarly discourse. This has given rise to the concern that Open Access “can become a tool of neocolonialism if it only gives students and academics better access to science from the North.”\textsuperscript{35} Indeed, studies show that gold Open Access has an exclusionary effect for independent scholars and scholars from less affluent institutions.\textsuperscript{36} Rather than as liberating tools, digital technologies thus risk to widen existing gaps and reinforce the position of hegemonic actors. In other words, “as a business model of academic capitalism, OA is already a reality; as a programme for bringing together the human race in intellectual dialogue and a common quest for knowledge, it remains a utopia.”\textsuperscript{37}

These trends and dynamics are even further accelerated in recent years. As part of a “massive adaptation process” in the digital era,\textsuperscript{38} some of the leading academic publishers have started to change their business models from providing academic content to becoming data analytics specialists. These developments are described in detail in the much-cited 2019 landscape analysis of the Scholarly Publishing and Academic Resources Coalition (SPARC), commissioned in response to the “growing trend of commercial acquisition of critical infrastructure” and co-written by a market analyst with long experience in the academic publishing market.\textsuperscript{39} With an increasing number of voices warning against this development,\textsuperscript{40} more recently science policy makers also started to take note of the developments.\textsuperscript{41}

In a nutshell, some of the leading academic publishers have started
to use the data tracking tools developed by the tech giants to collect and monetize data. The reason is not so much the direct monetary value of these data, but rather their value for predictive purposes.\textsuperscript{42} In other words, the data allow the publishers to improve and expand their services. In the digital era, the academic publishers do not limit themselves anymore to the dissemination of research output in the form of journal articles and books, but rather offer services covering an increasing range of the research (and educational) spectrum. By way of example, Elsevier calls itself a “global leader in information and analytics,”\textsuperscript{43} and Taylor & Francis describes its services as “content and research platforms.”\textsuperscript{44} The services offered today cover research assessment systems, productivity tools, and online learning management systems. In other words, publishers are gradually moving into the governance of the university and research institutes more broadly. By way of example, Elsevier’s “Pure” Research Management System “facilitates an evidence-based approach to your institutions’ research and collaboration strategies, assessment exercises and day-to-day business decisions.”\textsuperscript{45} In practice, the use of such services often forms part of packages or “big deals.”\textsuperscript{46} An example that caused an outcry in the scholarly community was a contract concluded between a consortium of Dutch Universities with Elsevier. While the deal foresees access to journals and Open Access publishing, in exchange it obliges the universities to license a large set of Elsevier’s products (including “Pure”).\textsuperscript{47} This is thus a new way (after big journal subscription deals) of selling unrelated services in packages—a strategy that is considered to

\textsuperscript{42} There is, however, evidence that in the US data have been sold by publisher to law enforcement authorities. See Sarah Lamdan, \textit{Librarianship at the Crossroads of ICE Surveillance}, \textit{In the Library with the Lead Pipe} (Nov. 13, 2019), http://www.inthelibrarywiththeleadpipe.org/2019/ice-surveillance/#identifier_4_9687.

\textsuperscript{43} See \textit{This is Elsevier}, ELSIEV (Feb. 26, 2022), https://www.elsevier.com/about/this-is-elsevier. (“As a global leader in information and analytics, Elsevier helps researchers and healthcare professionals advance science and improve health outcomes for the benefit of society.”).

\textsuperscript{44} See TAYLOR & FRANCIS, https://taylorandfrancis.com (last visited Apr. 15, 2023) (“Taylor & Francis offers a range of content platforms to connect readers to knowledge.”).


\textsuperscript{47} Claudio Aspesi, \textit{Leaked Dutch Contract with Elsevier Raises Significant Alarm Bells}, SPARC (Dec. 10, 2019), https://sparcopen.org/news/2019/leaked-dutch-contract-with-elsevier-raises-significant-alarm-bells/ (“According to press reports on leaked details of the deal, Elsevier is discussing a contract to provide Dutch universities with access to its journals at no extra cost (a major concession after decades of significant annual increases for most of their customers.”)
be advantageous for publishers and much less so for universities. As a result, publishers render themselves increasingly “indispensable for the governance of academic institutions and universities.”\textsuperscript{48} Large parts of the research and university infrastructure, the backbone of research and teaching, might soon be in the hands of commercial actors. This leads to a further market concentration, the loss of diversity and the building of monopolies or quasi-monopolies.\textsuperscript{49}

In conclusion, the internet, different than expected, did not mean the end of for-profit academic publishing—to the contrary, digital technologies allowed business actors to expand their power in the publishing sector. The data they collect from the users of their ever-more encompassing services offer insights into entire research cycles.\textsuperscript{50} Moreover, it gives these actors ample opportunities to “invisibly and strategically influence, and perhaps exert control, over key university decisions—ranging from student assessment to research integrity to financial planning.”\textsuperscript{51} This heavily increases concerns about the distortion of scientific processes by external factors. In other words, the privatization and commercialization of big parts of the digital research infrastructure in unprecedented ways risks infusing an economic rationale into the science system—with potentially corrupting effects.\textsuperscript{52}

B. Academics and the Scholarly Community: The Problem of Self-Corruption

However, academic publishers are only one part of the picture. The central role played by leading publishers can only be explained and understood within the broader context and dynamics of the science system. Indeed, publishers are so successful because of the strong demand on the other end. The current reward system in academia with reputation as its currency strongly incentivizes academics to publish their research in the most renowned outlets, which typically are owned

\begin{itemize}
  \item \textsuperscript{48} GER. RSCH. FOUND., supra note 42, at 6.
  \item \textsuperscript{49} ASPESI ET AL., supra note 47, at 21.
  \item \textsuperscript{50} GER. RSCH. FOUND., supra note 42, at 6.
  \item \textsuperscript{51} ASPESI ET AL., supra note 14, at 5.
  \item \textsuperscript{52} See Taubert & Weingart, supra note 3, at 10 (discussing structural developments leading to these concerns); Isabell Hensel & Gunther Teubner, \textit{Horizontal Fundamental Rights as Conflict of Law Rules: How Transnational Pharmagroups Manipulate Scientific Publications}, in \textit{CONTESTED REGIME COLLISIONS: NORM FRAGMENTATION IN WORLD SOCIETY} 139, 141 (Kerstin Blome et al. eds., 2016) (discussing publication bias resulting from conflict of interest in privately sponsored research on pharmaceuticals).
\end{itemize}
by leading publishers.53 This leads to the paradoxical situation that even though there is broad agreement in the scholarly community that some of the practices of leading publishers are highly problematic, at the same time scholars and academics play the main role in keeping the system alive.54 Part of the explanation for why the transition to Open Access proceeded much slower than expected is also that scholars hesitate to accept this new publication mode, and at times even resist it, as a case currently pending before the German Constitutional Court evidences.55 When the University of Konstanz introduced a regulation obliging its members to publish their research Open Access—under certain circumstances and after an embargo period—members of the law faculty saw their academic freedom violated, and the faculty attacked the regulation before the competent courts.57

The importance of publications is closely connected to the role and function of reputation in the science system. On the one hand, publications build the basis for the scientific endeavor, since only published results of research can be tested and refuted or further advanced. On the other hand, the publication has an important social role in the science system, adding to the reputation of the respective researcher— and determining future career paths.58 Unsurprisingly, the publication process in science from submission to publication and dissemination is strongly formalized and regulated, having a strong disciplining effect, and leading to deep path-dependencies. In the literature, it is recognized that the relevant actors shaping the publication process next to publishers and researchers are also


54. An example is the scholar-led call to boycott the publisher Elsevier especially because of the “exorbitantly high prices for subscriptions to individual journals.” Over 20,400 people from all disciplines have signed. See 20573 Researchers Taking a Stand, THE COST OF KNOWLEDGE, http://thecostofknowledge.com/.


57. The administrative court (Verwaltungsgerichtshof) of Baden-Württemberg decided on September 26, 2017 to suspend the proceeding and ask the Constitutional court for a preliminary ruling. See VGH Baden-Württemberg, Decision of 26 September 2017, 9 S 2056/16. The question before the Constitutional Court now, however, relates to the powers to adopt said regulation rather than one of academic freedom.

universities and their appointment and promotion procedures. The term “publication regime” has been proposed to describe the interlocking of these different aspects and actors. Younger scholars are “disciplined” into traditional paths if they want to make their way up in the system’s hierarchy. Options to deviate are therefore limited. This explains why, despite the deep changes publishing is undergoing in the course of the digital revolution, processes and formats remain surprisingly stable. This also explains why publication modalities, such as the decision where to publish, play such an important role. Publishers have become a surrogate for reputation, which in turn is a surrogate for truth. Against this backdrop, the success of the gold Open Access model is hardly surprising—it is the model that deviates least from the existing path, leaving the central role of traditional publishers untouched.

C. Science Policy Makers: An Increasingly Instrumental Reading of Science

It could thus be said that it is the expansionary tendencies of the economic system in combination with self-endangerment and the corruption potentials of academics that account for the dynamics in the science system. Yet, at a closer look an additional actor comes into play. Science policy makers increasingly get involved and play an active role in steering and shaping the developments in the science system. This is particularly evident with Open Access policies.

While the Open Access movement initially consisted of a rather small group of mainly activists, public actors at national and international levels increasingly got involved, the main argument being the large sums of public money spent. In 2019, UNESCO issued a recommendation on Open Science and in 2020, the UN Committee on Economic, Social and Cultural Rights addressed Open Science as part of the so-called right to science under Art. 15 of the International Covenant on Economic, Social and Cultural Rights. Today, most countries and also the European Union have official Open Access

59. Erik Cohen et al., The Global Permutations of the Western Publication Regime, 21 CURRENT ISSUES TOURISM 2035, 2035 (2018); Carlos Spoerhase, Filetierte Vernunft, 36 MITTELWEG 1, 2 (2022).
60. Teubner, supra note 55, at 512.
62. General Comment No. 25, supra note 28.
policies. The Swiss strategy, for example, envisages that, by the year 2024, Open Access should be the full reality,64 and the Swiss National Science Foundation “requires grantees to make the results of SNSF-funded projects available in an open access publication or database.”65 In addition, there exist numerous international initiatives and transnational networks such as Plan S by the so-called cOAlition S, supported by a diverse conglomerate of public and private research funders.66

While together these attempts have certainly contributed to the Open Access transformation, new criticism has arisen, mainly because of the prominent role many of these initiatives attribute to leading commercial publishers. The justification usually given is the preference of researchers to publish their work in the most renowned journals. An example is the controversial “Projekt DEAL” in Germany, whose focus on “publish and read” agreements with the leading publishers has led critics to argue that this strengthens and perpetuates the oligopolistic position of a few powerful publishers.67 Others argue that this goes to the detriment of other visions and concepts of Open Access, such as the non-commercial platforms prevalent in Latin America.68 In other words, the support of powerful and mainly Western science funders has made an important contribution in making gold Open Access the standard—despite its proven problematic and exclusionary effects.

Additionally, policy makers also shape the developments in more subtle, but no less important ways. Of particular relevance in the present context is the metrification of science as one aspect of the structural transformation science is currently undergoing.69 Output

---

69. On the structural changes, see Leonhard Dobusch & Maximilian Heimstädt, The Structural Transformation of the Scientific Public Sphere: Constitution and Consequences
measuring today takes place at essentially all levels, from the ranking of universities and journals to the assessment of the impact and performance of individual researchers. Early forms of these mechanisms already emerged in the mid-twentieth century from within the scientific community in the form of citation indexing. The aim was to support and facilitate orientation within a landscape marked by a growing number of publications by statistically understanding and showing publication and citation patterns. The digitalization massively supported and sped up this development. With the gradual neoliberal reform of the university system in most European states and the United States and the shift to new public management approaches in the governance of universities and research institutions, these instruments have been picked up by science policy makers (and in the meantime by publishers).

They have thus become a significant external source of influence and possible intervention both from the economic, but also the political system. This is particularly evident when quantifiable performance is coupled with financial advantages, be it in the form of higher salaries or more funds.

While digitally facilitated bibliometric indicators promise to translate information about research activities “into numbers that, in their apparent neutrality, seem to transcend linguistic and cultural (including disciplinary) boundaries,” for years they have been at the heart of critical debates and in the meantime have even produced counter-initiatives. On the one hand, the purported neutrality and objectivity has long been put into question, and the approximation of

of the Path Towards Open Access, 21 WEIZENBAUM SERIES 1 (2021); Taubert & Weingart, supra note 3.
70. For a detailed account, see YVES GINGRAS, BIBLIOMETRICS AND RESEARCH EVALUATION: USES AND ABUSES (2016).
72. Taubert & Weingart, supra note 3, at 11; on the role of publishers, see Jake Goldenfein & Daniel Griffin, Google Scholar – Platforming the Scholarly Economy, 11 INTERNET POLICY REV. 1 (2022); on the issue of possible conflict of interest between disseminating and evaluating research, see ASPESI ET AL., supra note 47, at 15–16.
73. For an explanation of the situation in Germany, see STEFAN HORNBORSTEL, WISSENSCHAFTSINDIKATOREN: BEWERTUNGEN IN DER WISSENSCHAFT 40 (1997); Max-Emanuel Geis, Eigengesetzlichkeit als Strukturprinzip der Wissenschaft, GLANZLICHTER DER WISSENSCHAFT 39 (2015).
74. Lynn P. Nygaard & Rocco Bellanova, Lost in Quantification: Scholars and the Politics of Bibliometrics, in GLOBAL ACADEMIC PUBLISHING. POLICIES, PERSPECTIVES AND PEDAGOGIES 23 (Theresa Lillis & Mary Jane Curry eds., 2018).
quality and quantity has been criticized. On the other hand, the quantification and metrification produced unintended side-effects. It introduced strong incentives to publish even more (“publish or perish”), leading academics to “salami slice” their results into as many publications as possible or even to favor more mainstream topics that promise to “earn” more citations. Open Access to some extent exacerbates this logic—studies show a link between the open accessibility and increased visibility through higher downloads and citation rates. Clearly, these seemingly purely infrastructural decisions thus exert an influence on research design, and, somewhat paradoxically, they have further exacerbated the market logic in science. Commercial publishers have long discovered the business potential and offer a number of digital tools for evaluative metrics and analytics as next generation products. We have thus come full circle.

III. UPDATING THE CONSTITUTIONAL PROTECTION OF SCIENCE IN THE DIGITAL AGE

The internet and digital tools offer important chances for science. However, as seen, they also enhance certain problematic tendencies in the science system, ultimately contributing to further strengthening the powerful position of a small group of commercial publishers. While the internet and digitality are not the reason for these developments, they thus exacerbate, accelerate, and magnify pre-existing tendencies. The aim of this section is to explore the constitutional protection of the science system against the backdrop of these threats. It is argued that while traditional accounts of constitutionalism are not entirely blind to pressures for the autonomy and integrity of science beyond the nation-state, a reading of academic freedom informed by societal constitutionalism helps to better grasp and address the dangers science faces in the digital era.

A. The Institutional Dimension of Academic Freedom in Constitutional Law

Also from the perspective of traditional state-centered

76. Nygaard & Bellanova, supra note 74, at 24. For a detailed account, see Gingras, supra note 70.

77. For a more differentiated finding, see James M. Donovan & Carol A. Watson, Citation Advantage of Open Access Legal Scholarship, 103 L. LIBR. J. (2011); Allison Langham-Putrow et al., Is the Open Access Citation Advantage Real? A Systematic Review of the Citation of Open Access and Subscription-Based Articles, 16 PLOS ONE (2021).

78. For an analysis of Google Scholar, see Goldenfein & Griffin, supra note 72.
constitutionalism, the above-described developments in the science system are seen as potential threats to fundamental rights. On the one hand, the described tracking and surveillance activities may amount to infringements of the academic freedom and privacy rights of individual researchers. Indeed, these digital tools open wide possibilities for governments to surveil researchers under the guise of national security or public order, possibly leading to self-censorship. In the United States, instances of publishers selling sensitive data to law enforcement authorities have furthermore been reported. Another example of possible rights violations in this context is provided by the Universities of Liverpool and Leicester which, based on metrics, planned major staff cuts in the course of the COVID-19 pandemic.

On the other hand, the above-described developments also threaten the autonomy and integrity of science as a system or institution. As shown above, they facilitate corporate control and influence of science, leading to an increasing intrusion of an economic rationale into the science system. At the same time, they increase self-destructive tendencies within the science system which in turn are fed by demands to measure and quantify research output.

Constitutional law de lege lata protects the autonomy of science, at least in parts. By way of example, German constitutional doctrine has long established that the fundamental right of academic freedom as enshrined in Article 5(3) of the Basic Law not only shields individuals from state interference, but also obliges the state to provide a “positive order for science,” thus recognizing an “objective” or institutional dimension of academic freedom. This does not mean that the state should actively regulate science—very much to the contrary, the idea is that the state should create the conditions under which science can develop and unfold according to its own rationale (Eigengesetzlichkeit in German). The protection accorded to science has thus aptly been

80. See Lamdan, supra note 43.
82. Ulrich Karpen, Die Finanzverfassung der Hochschulen Angesichts der Gegenwärtigen Haushaltslage, 73 VERWALTUNGSARCHIV 405, 406 (1982). For the difference between the “subjective and objective” dimensions of academic freedom, see Ino Augsberg, Subjektive und objektive Dimensionen der Wissenschaftsfreiheit, in FREIHEIT DER WISSENSCHAFT. BEITRÄGE ZU IHRER BEDEUTUNG, NORMATIVITÄT UND FUNKTION 60 (Friedmann Voigt ed., 2012).
83. On the difference between the subjective and objective dimension of academic freedom, see Augsberg, supra note 82.
described as “in between fundamental freedom and state institutionalization”. Ultimately, the protection of the autonomy of science is based on the premise that under these conditions, science thrives. According to the German Federal Constitutional Court, freedom of science is “based on the idea that a science free of social utilitarianism and political expediency best serves the state and society as a result.” In other words, the idea is to ensure “those minimum conditions which a free science must fulfil in order not to undermine the function of its protection of freedom.”

At the international level as well, a certain institutional protection of science is recognized. The recent General Comment No. 25 on the so-called right to science is worth mentioning in this context. In this General Comment, the Committee on Economic, Social and Cultural Rights highlighted that positive obligations play a particularly important role when it comes to science. Recognizing the international nature of science, the Committee furthermore made it clear that the “right to science” requires states “to take steps through legislation and policies, including diplomatic and foreign relations, to promote an enabling global environment for the advancement of science . . . .”

Constitutional scholarship has also discussed threats to science from the private sphere. While protection historically was directed against religious and political interference, today, the role of private actors in, and the economization of, science is widely discussed. This concerns in particular collaborations between science and industry and private research funding, but also issues of publication bias. Also the metrification of science and an increasing competition mentality, said to

---

84. HANS-HEINRICH TRUTE, DIE FORSCHUNG ZWISCHEN GRUNDRECHTLICHER FREIHEIT UND STAATLICHER INSTITUTIONALISIERUNG: DAS WISSENSCHAFTSRECHT ALS RECHT KOPERATIVER VERWALTUNGSVORGÄNGE (1994).
85. Bundesverfassungsgericht [BVerfG] [Federal Constitutional Court] July 14, 2022, docket number 1 BvR 748/06 (Ger.), http://www.bverfg.de/e/rs20100720_1bvr074806.html (translation by the author).
86. KLAUS F. GÄRDITZ, UNIVERSITÄRE INDUSTRIEKOOPERATION, INFORMATIONSZUGANG UND FREIHEIT DER WISSENSCHAFT 75 (2019) (translation by the author).
88. Id. ¶ 77 (emphasis added).
89. See GÄRDITZ, supra note 86. See generally ISABEL HENSEL, GRUNDSRECHTKONFLIKTE IM VERTRAG: VERTRAGSVERFAHRENRECHT FÜR DIE KOPROPTION VON HOCHSCHULEN UND WIRTSCHAFT (2021) (discussing how the conflict between universities and private actors can lead to problems for science).
increase non-scientific incentives in the science system, has been discussed as a problematic instance of interference with science. In this context, the goal of academic freedom is to protect scientists against “commercialization pressure” and to “sufficiently contain the dangers this poses for a neutral, detached science oriented towards “finding the truth.” To put it somewhat bluntly, academic freedom also requires protecting the “scientific process . . . from contamination that is inadequate to science.” The aim is thus also to protect the credibility and reliability of science as such, or its “independence.” It is also worth mentioning that the Committee on Economic, Social and Cultural Rights in the above-mentioned General Comment No. 25 discussed the complex relationship between science and intellectual property rights as one instance of commercialization or privatization of science, addressing among other things the “excessive price of some scientific publications.” Yet, the Committee refrained from deducing concrete legal consequences from this finding.

B. The Added Value of Societal Constitutionalism

As just explained, the doctrine on the “objective” dimension of academic freedom shows that more classic accounts of constitutionalism are not entirely blind to constitutional challenges transcending the state-individual relationship. Yet, the shortcoming of the classic approach is that it ultimately relies only on state and political power to solve societal issues, shifting the responsibility entirely to public actors, which is particularly inadequate in a globalized setting. Furthermore, there is the risk that more subtle forms of threats, such as those

91. For German constitutional doctrine, see Christian Bumke, *Universitäten im Wettbewerb*, 69 VERÖFFENTLICHUNGEN DER VEREINIGUNG DER DEUTSCHEN STAATSRETSLEHRER 462 (2010); Geis, supra note 73.
92. GÄRDITZ, supra note 86, at 81 (translation by the author).
93. Id. at 74 (translation by the author).
94. Id. at 78 (translation by the author).
96. General Comment No. 25, supra note 28, ¶ 61.
TACKLING THREATS TO ACADEMIC FREEDOM BEYOND THE STATE 283

provided by data tracking and the “privatization” of research infrastructures, which often will remain under threshold of specific and manifest manipulations and processes, remain outside of the scope.99

A reading of academic freedom inspired by societal constitutionalism would help to capture these more subtle threats arising in the digital context and falling outside of the traditional state-individual schema. This would not require the adoption of a new right but could be achieved through an updated interpretation of academic freedom.100 The recent General Comment No. 25, in which the UN Committee on Economic, Social and Cultural Rights interpreted Open Access as one aspect of the “right to science,”101 is one example of many showing how existing rights can be interpreted in light of new situations. One of the core messages of societal constitutionalism is that the function of fundamental rights in modern societies should not only be to protect individuals against excesses by political power, which is at the heart of traditional state-centered constitutionalism. Instead, it argues that fundamental rights also have a broader, collective function, protecting functional sub-systems from their encroachment by other systems, including the economic system.102 Indeed, societal constitutionalism highlights the “trans-subjective potential of subjective rights.”103 In other words, fundamental rights also play a role in ensuring social differentiation and the mutual co-existence of societal sub-systems. Taking such a reading into account when interpreting the institutional dimension of academic freedom could thus help to protect the science system from subtle, yet systemic threats, such as those posed by data tracking and the large-scale acquisition of research

100. On different strategies to cope with the challenges for human rights posed by digital technologies, see Dafna Dror-Shpoliansky & Yuval Shany, It's the End of the (Offline) World as We Know It: From Human Rights to Digital Human Rights – A Proposed Typology, 32 EUR. J. INT. LAW 1249 (2021).
101. See General Comment No. 25, supra note 28, ¶ 1.
infrastructure by private actors, provided one agrees that these practices threaten the autonomy of science. In conclusion, societal constitutionalism can provide useful insights for the discussion about digital constitutionalism and the effective protection of fundamental rights such as academic freedom in the digital era.104

IV. REGULATORY RESPONSES TO THE DIGITAL THREATS FOR SCIENCE

The previous sections have shown that societal constitutionalism provides a useful lens for better understanding and conceptualizing threats to fundamental rights arising in a digital context. Yet, also for the question of how to respond to these constitutional challenges, the theory has something to say. The basic assumption of systems theory is the idea of autopoiesis (i.e., the operational closedness of systems). From this follows that systems cannot directly influence each other, or in other words, information can never pass directly from one system to another. However, this does not mean that systems are sealed off from their environment—to the contrary, they are “cognitively open” and interact through structural couplings, understood as situations in which “a system presupposes certain features of its environment on an ongoing basis and relies on them structurally . . . .”105 Information received from the outside thereby is translated into the code of the system—possibly triggering “irritations, surprises, disappointments, disturbances.”106 Taking this as a starting point, societal constitutionalism has developed useful insights about the role of law in steering change and development in other societal systems. It argues that for legal regulation to be effective, or in other words, to produce the intended outcome, it needs to take into account the inner rationale and logic of a given system (“reflexive law”).107

The aim of this last part of the paper is to offer preliminary thoughts about possible legal and regulatory responses to the challenges the science system currently faces. Engaging with societal constitutionalism as a normative theory and using its theoretical insights about the role of law in addressing structural challenges, it argues that, on the one hand, self-constitutionalization processes in the

105. E.g., NIKLAS LUHMANN, DAS RECHT DER GESELLSCHAFT 441 (1993).
106. Cf. id. at 442; NIKLAS LUHMANN, DIE WISSENSCHAFT DER GESELLSCHAFT 165 (1990).
107. GUNTHER TEUBNER, LAW AS AN AUTOPOIETIC SYSTEM 64-99 (1993).
science system should be strengthened. On the other hand, strong limitative rules to contain the superimposition of power and profit over science are needed.

A. Supporting the Constitutionalization of the Science System

It is often said that science has come to its autonomy on its own (i.e., with little support by law). Indeed, the science system has traditionally been considered a system where forms of self-governance prevail over state regulation. In other words, “science basically has arrived at its autonomy without legal support.”\(^\text{108}\) The legal system explicitly recognizes its own limits to regulate science. As described above, constitutional doctrine in Germany starts from the idea that it is the state’s role to provide the conditions and, indeed, create the space under which science can develop according to its inner logic or \textit{Eigengesetzlichkeit}.\(^\text{109}\) Also, the Committee on Economic, Social and Cultural Rights in the already mentioned General Comment No. 25 has stressed the limits of law as a tool to steer science, in this concrete case towards more open research practices:

\begin{quote}
[O]pen science cannot be achieved by the State alone. It is a common endeavor to which all other stakeholders should contribute, nationally and internationally, including scientists, universities, publishers, scientific associations, funding agencies, libraries, the media, and non-governmental institutions. All these stakeholders play a decisive role in the dissemination of knowledge, especially when it comes to outcomes of research financed with public funds.\(^\text{110}\)
\end{quote}

However, this does not mean that there is no role for external legal intervention in processes of self-regulation and self-constitutionalization. On the contrary, societal constitutionalism explicitly acknowledges a supplementing and supporting role of state law in stabilizing these inner processes. In this sense, the need for external intervention is not a theoretical incoherence of the theory; rather, the argument is that if external intervention happens to support functional differentiation, it should respond to the inner rationality of

\begin{footnotes}
\item[108] TEUBNER, supra note 18, at 108.
\item[109] See supra section III.A.
\item[110] General Comment No. 25, supra note 28, ¶ 49.
\end{footnotes}
each system and, indeed, can only be effective in its task if it does. In other words, societal constitutionalism does not claim that “particular global social spheres such as the internet, the economy, or the field of health care can provide enough counterweight when state constitutional norms are lacking.”

As shown above, actors from within the science system have started reacting to the new challenges, or, in other words, self-regulatory processes are already underway. An increasing number of reports are appearing, and initiatives are emerging against the use of data tracking tools in science. Furthermore, an alternative Open Access model has developed and is increasingly gaining ground, namely the so-called diamond or platinum model. While there exists slightly different definitions of diamond Open Access, their common denominator is that fees are neither charged for authors nor for readers, in contrast to the gold model and subscription-based journals. More recent initiatives highlight that the diamond model should be combined with community-driven governance of these journals. In other words, not only should publishing be free of charge, but journals should also be governed and managed from within the scholarly community itself. Also the reformulated version of the Budapest Open Access Initiative recommends to “move away from Article Processing Charges” and endorses “community-controlled infrastructure.” Similar to other digital commons such as Wikipedia, the diamond model thus offers an alternative to either public or private provision, placing the dissemination of research “beyond market and state.” Yet, the diamond model suffers from two major limitations: diamond Open Access journals often lack sustainable funding, and newly funded journals struggle to become true alternatives to their already reputable

111. Golia & Teubner, supra note 19, at 388.
117. The Budapest Open Access Initiative, supra note 15.
commercial counterparts.\textsuperscript{119}

External intervention can thus support the ongoing self-regulatory processes in at least two ways. First, existing funding streams should be diversified and redirected to include Open Access models such as diamond Open Access. This might include reforms of budgetary rules and the organization of libraries. Existing Open Access funding strategies regularly favor gold Open Access. A case in point is the German “Projekt DEAL”, which exclusively focuses on big deals with commercial publishers as a strategy to contribute towards the Open Access transformation.\textsuperscript{120} While the idea is to mimic the science system’s rationale by following the preferences of researchers, the result has been to reinforce certain problematic tendencies and above all to strengthen and perpetuate the oligopoly of leading publishers.

Second, to become a true alternative, publishing formats and models need to find acceptance within the scholarly community. In light of the deeply entrenched path dependencies, the establishment of accepted new formats in academia is not straightforward. As shown above,\textsuperscript{121} the incentives in place and the existing assessment and reward system support problematic publish-or-perish tendencies, which ultimately serve the big publishers. Rather than being absent in this development, the state, through neoliberal reforms of the university system and competitive funding structures, actively contributes to them. Ultimately, a true path change can only be achieved if the underlying system of incentives is tackled. In this spirit, the 2022 Budapest Declaration recommends reforming research assessment: “Careful reforms can eliminate disincentives for OA and create new incentives for OA.”\textsuperscript{122} To reach this goal, research funders should support initiatives such as the Declaration on Research Assessment (DORA)\textsuperscript{123} or the Leiden Manifesto,\textsuperscript{124} which aim to improve the evaluation system in academia. Both documents problematize the widespread use of metrics and the “journal impact factor” as a means to evaluate research which, among other problems, creates incentives to publish with the most impactful journals. While the DORA declaration recommends the elimination of journal-based metrics altogether, the Leiden Manifesto

\textsuperscript{119. See generally Stephanie Normand, Is Diamond Open Access the Future of Open Access?, 3 THE IJOURNAL (2018).}
\textsuperscript{120. Supra section II.C.}
\textsuperscript{121. Id.}
\textsuperscript{122. The Budapest Open Access Initiative, supra note 15, at 7–9.}
\textsuperscript{123. The Declaration on Research Assessment, developed in 2012 at the Annual Meeting of the American Society for Cell Biology in San Francisco. DORA, https://sfdora.org.}
\textsuperscript{124. Diana Hicks et al., Bibliometrics: The Leiden Manifesto for Research Metrics, 520 NATURE (2015).}
demands that quantitative evaluation should be supported and complemented by qualitative means. However, it remains questionable whether such initiatives can be successful and are indeed compatible with the competitive public funding schemes in place in many countries, which feed an economic, output-oriented rationale in the science system.

B. Limiting Excesses of the Economic System

Besides the support of processes internal to the science system, externally, the influence of commercial actors with expansive and colonizing tendencies needs to be limited.\textsuperscript{125} Indeed, in light of the fact that it is unlikely that actors of the digital economy will adopt limitative rules voluntarily, business models built on "surveillance capitalism"\textsuperscript{126} call for a strong intervention by states and political actors.\textsuperscript{127} Taming the power of big tech in the digital age is a hotly debated issue, and the EU is taking the role as a global forerunner in this respect. In light of the transnational nature of these challenges, this orchestrated action seems to provide a useful starting point, and the recently adopted instruments, as well as those still in the pipe, will certainly continue to give rise to vivid discussions in the years to come. From the perspective of societal constitutionalism, the core question will be whether these instruments sufficiently manage to respond to the dynamics of the economic system in the digital age. To give a concise overview of these attempts and their application to the case at hand deserves an in-depth analysis which would exceed the scope of this paper, but a short preliminary overview is in order.

Relevant for the present context are, on the one hand, discussions about increasing the transparency of algorithms which, as shown above, are also used in science, including tools to evaluate and rank scholars, such as Google Scholar.\textsuperscript{128} The "complex and black-boxed architecture"\textsuperscript{129} of big tech and their opaque use of algorithms, leading

\begin{itemize}
\item \textsuperscript{125} Lars Viellechner, The Transnational Dimension of Constitutional Rights: Framing and Taming "Private" Governance Beyond the State, 8 GLOB. CONSTITUTIONALISM 639, 640 (2019).
\item \textsuperscript{128} Goldenfein & Griffin, supra note 72, at 10.
\item \textsuperscript{129} Ganaele Langlois et al., Networked Publics: The Double Articulation of Code and Politics on Facebook, 34 CAN. J. COMM'C\textsuperscript{N} 415, 416 (2009).
\end{itemize}
to “informed realities,” has long been recognized as a problem. Regulatory attempts to increase the transparency and accountability of algorithms include the “right to explanation” in the EU’s General Data Protection Regulation, which, despite numerous open questions, according to commentators are going in the right direction. Also the recently enacted EU Digital Services Act (DSA) and the proposed EU AI Act call for more transparency.

On the other hand, it is well recognized that with its expansive tendencies, the economic system risks harming itself. In the digital context, the practice of self-preferencing is an example of compulsive growth tendencies. Resorting to antitrust instruments seems an obvious solution, and it is therefore a welcome step that the EU has complemented its digital strategy with the Digital Markets Act (DMA), aiming to ensure fair competition online. However, with its rather narrow focus on “gatekeepers,” i.e., actors that have already reached a status in which they control markets, the DMA will most likely not be applicable to academic publishers. Legislators have not yet recognized the dangers of “surveillance capitalism” for science.

V. CONCLUSION

For a long time, the science system offered a textbook example of self-regulation or “regulation beyond the state.” It was considered a system that needed little outside support to reach and maintain its autonomy. Yet today, it also offers a textbook example of constitutional challenges and threats related to digitality. In the early years of the internet, there was great enthusiasm about the nearly infinite possibilities this new medium offered for the distribution and accessibility of information. In the meantime, however, awareness about the darker sides of digital technologies is on the rise since it has become known that the leading academic publishers have long started engaging

131. See generally Jørgensen, supra note 16.
135. See, e.g. DMA, supra note 10, ¶ 52.
136. Id.
in the data analytics business. While commercialization and marketization as a threat to the autonomy and integrity of science have been discussed for a long time, the growing practice of data tracking and the resulting datafication strongly increases concerns about corporate capture and the platformization of science. Digital technologies, while holding a wealth of new possibilities for science, at the same time have created new means of value extraction and commodification. In other words, they accelerate and enhance the expansive tendencies of communicative media, such as the commodification imperative of the economy. This threatens not only privacy rights and the freedom of individual academics and researchers but also the autonomy of science as an institution or system.

Is constitutional law prepared to address these threats and challenges of the digital age? This question is of utmost importance in light of the still-growing power shift to private actors and the totalizing tendencies we are witnessing today. The aim of this paper was to explore the constitutional protection of science and discuss the potential of societal constitutionalism in filling protective gaps in the digital age. While the consequences of the increasing corporate control of media and information infrastructure for free speech and the public sphere are extensively discussed today, this is much less the case for the science system. The analysis has shown that while classic accounts of constitutionalism are not entirely blind to threats by private actors, the subtle yet systemic risks the science system faces in the digital age remain under the threshold of specific and manifest manipulations and processes.

Which lessons can we thus draw from societal constitutionalism for the protection of rights against the threats arising in the digital age? Highlighting the institutional and collective dimension of rights, one of the core messages of societal constitutionalism is that the function of fundamental and human rights in modern societies should not only be to protect individuals against excesses by state power but also societal sub-systems from their mutual encroachment, upholding social differentiation, and ensuring the co-existence of societal sub-systems. Taking such an institutional reading into account when interpreting the “objective” dimension of academic freedom could thus help to protect the science system from threats such as those posed by data tracking and the large-scale acquisition of research infrastructure by private actors.

Yet, societal constitutionalism not only helps to better conceptualize and grasp constitutional threats in relation to digitality but also to address them. Starting from the idea that law needs to consider the inner logic of a given societal sub-system to be effective, societal constitutionalism also provides answers to the very question about the
role law can play in tackling present-day challenges. Based on the theory of reflexive law as promoted by societal constitutionalism, this paper has thus also addressed the question of how self-constitutionalization processes in the science system can be supported “from the outside” and how limitative rules may contain the superimposition of power and profit over science, creating and protecting space for digital infrastructure that follows neither a market nor a state logic.

In conclusion, the analysis of the current developments in science has shown that societal constitutionalism provides useful and multifaceted insights for the discussion about digital constitutionalism and the effective protection of fundamental rights in the digital era, questions that touch upon the very role constitutional law can play in offering protection against present-day threats that will certainly not lose any of their currency in the years to come.
Robo Justice: Constitutional Issues with Judge AI

TANIA SOURDIN*

ABSTRACT

The emergence of forms of Artificial Intelligence (AI) to support human judges or potentially replace them raises questions concerning the role of judges in contemporary society, as well as broader questions linked to the ongoing role and nature of the so-called “digital Switzerland” and their role in justice systems. In addition, tensions between the executive and the judicial branches of government with respect to the development of Judge AI support an exploration of societal constitutional analysis which requires consideration that reaches beyond mere institutional analysis. The varying and developing relationships between judges, courts, and emergent technologies challenge conventional theoretical approaches to governance and also require a greater focus on social interaction to explore how adaptation in an era of judicial responsiveness might support the development of ethical approaches that respond to vulnerable populations and social needs. Using an approach based on societal constitutionalism enables the development of a contemporary approach to justice that has been extended with AI initiatives. This involves setting limits to resist authoritarianism by framing the definition of “justice” so that it includes an overarching focus on promoting human wellbeing which in turn promotes restraint and thoughtful approaches toward the use of more disruptive technologies in the justice sector. The challenges presented by this approach in a global context are readily apparent in the context of differing justice conceptions and a focus on “fast,” “low cost” justice delivery, perhaps in the absence of justice itself.

Keywords. Judging, Artificial Intelligence, Social Interaction, Judges, Justice Definition.
I. Introduction

The judiciary plays a critical role in supporting societies that abide by the rule of law. This role is even more relevant when a society adopts a democratic governance structure. In many Western states, they form one of the three arms of government in conjunction with the executive and legislature. Neither the rule of law nor the judiciary is static and the 21st century’s social dynamics, particularly visible in the context of technological changes, have and will continue to shape the justice system with potentially “productive and destructive” effects. In particular, the reshaping of the justice sector as a result of technological changes has implications for the status and role of the judicial arm of government and requires a rethinking of constitutional arrangements.

This article focuses on the emergence of Judge AI and its impact on judicial functions and the delivery of justice. In discussing the development of Judge AI, the role of the judge and other actors in contemporary society, the challenges raised by such technological shifts to conventional theoretical approaches are explored. Suggestions are offered for the creation of a constitutional framework that facilitates justice through re-framing the definition of “justice” so that it includes an overarching focus on human wellbeing to promote restraint and thoughtful approaches towards the use of more disruptive technologies in the justice sector and advocating for judge-led reform through an ethical framework. By retaining a focus on human judge-led reform, some of the serious concerns that arise in the context of Judge AI and supportive Judge AI can be addressed.

While this article does not explore the potential issues that are raised by neuro tech, human augmentation and the development of “enhanced” judges, a “Judge v Robot” distinction has been employed to enable analysis and discussion. The response to such developments

---

*Tania Sourdin, Dean and Professor of Law, Newcastle Law School, University of Newcastle, Australia. The author thanks and acknowledges the kind assistance of researchers Charlotte Kuszleyk and Kate Slowey in the preparation of this article. The author also acknowledges that parts of this article draw upon TANIA SOURDIN, JUDGES, TECHNOLOGY AND ARTIFICIAL INTELLIGENCE (2021).


2. See Allan McKay, Neurorights: The Chilean Constitutional Change, AI & Soc (2022) (discussing the idea that in reality, the distinction is not binary in that it seems probable that developments in supportive Judge AI as well as human augmentation will challenge constitutional frameworks in additional and differing ways. For a discussion more focused on neuro technology, human augmentation and constitutional frameworks including the new constitutional framework being considered in Chile, a Chilean constitutional response to emerging neurotechnology).
may include consideration of differing human rights frameworks that incorporate rights such as the right to equal access to mental augmentation or the right to protection from algorithmic bias. Ultimately, such additional human rights framework approaches will also have an impact on constitutional reform; however, this article focuses more on activities and reform measures that could be specifically directed at the judicial role.

Some general risks have been articulated in the context of AI ethical literature and in regulatory instruments that include the proposal developed in Europe in 2021 for an Artificial Intelligence Act (AI Act). The proposed AI Act offers a more defined framework through proposed legislation, aimed to regulate “high risk” AI and make illegal AI that poses too hazardous an outcome while enabling investment and innovation in AI and promoting international cooperation. Moreina and Borges have praised the AI Act for identifying AI systems used in legal applications and justice administration as “high risk.” High risk provides that the AI approach must conform to a number of regulations. However, arguably the AI Act focuses on the individual rather than societal harm, with commentators noting that “...while many risks from AI technology can be thought of as the potential for harms to an individual’s health and safety or of adverse impact on their fundamental rights, AI may also cause significant harm, on a societal level.”

In this article, the potential impacts are explored by reference to ethical approaches noting that regulatory approaches may be


6. Id.


8. Id. at 6.

inadequate to attend to a broader range of concern areas. Initial discussion focuses on the emergence of Judge AI before considering the role of judges and the rise of technological giants. From this perspective, ethical frameworks are analysed with reference to the broader role judges play within society (beyond a determinative function) to support a process approach with regard to societal constitutionalism to be formulated.\(^{10}\)

II. THE EMERGENCE OF JUDGE AI

The question is not if artificial intelligence will reshape judicial functions and the justice sector but when.\(^{11}\) The legal profession's past reticence to technology adoption has changed dramatically with the gradual development of online courts, the digitization of court records, and the increasing use of AI to make or assist in making decisions outside the courtroom.\(^{12}\) At the most basic level, technology can assist in informing, supporting, and advising people involved in justice activities (“supportive technologies”). At the second level, technology can replace activities and functions that were previously carried out by humans (“replacement technologies”). Finally, at a third level, technology can provide for very different forms of justice, particularly where processes change significantly (“disruptive technologies”).\(^{13}\) It is at the third “disruptive” level that technology can “displace and revolutionize conventional working habits and bring radical change.”\(^{14}\) The technological developments which have already taken place in the legal profession make it more likely that forms of AI will be used to support and possibly supplant some judicial decision making in the future.\(^{15}\)

In a practical sense, AI has been defined as “the theory and development of computer systems able to perform tasks that normally


\(^{13}\) See generally Tania Sourdin, *Justice and Technological Innovation*, 25 JOURNAL J.A. 96, 105 (2015) (material generally drawn from these sources); see also Tania Sourdin et al., *Just, Quick and Cheap? Civil Dispute Resolution and Technology*, 19 MACQUARIE L.J. 17, 19 (2019) (discussing the same taxonomy); Sourdin, supra note 11, at 1114.


\(^{15}\) Sourdin, *supra* note 11, at 1117.
require human intelligence.”16 In a technical sense, however, AI is related to machine learning, natural language processing, expert systems, vision, speech, planning, and robotics.17 By incorporating both meanings, this article uses AI as an umbrella term that encompasses branches of science and technology, which often involves the creation of complex algorithms to enable decisions to be made.18

Judge AI refers to developments in the various branches of AI specifically concerned with contributing to judicial tasks and can incorporate a range of disruptive technologies.19 It includes a range of possibilities: from the increasing use of technology in administrative and judicial processes prior to trial20 to playing a role in decision making processes (“supportive Judge AI”) or replacing a judge altogether (“Judge AI”).21 Examples of supportive Judge AI already in place include the production of template decisions which judges use as a base for explaining their reasons for judgment.22 More sophisticated forms of Judge AI are also being used to “nudge” or “correct” judicial decision making in countries such as China.23 Supportive Judge AI can raise distinct issues relating to judicial independence, and although such issues may not raise constitutional issues, it is possible that they may do so if they are perceived as having an impact on judicial independence. As previously noted, “[s]upportive Judge AI arrangements could also be perceived to be problematic in some countries, particularly if such supports are regarded as having an

16. DAVID SCHATSKY, CRAIG MURASKIN & RAGU GURUMURTHY, DEMYSTIFYING ARTIFICIAL INTELLIGENCE: WHAT BUSINESS LEADERS NEED TO KNOW ABOUT COGNITIVE TECHNOLOGIES 3 (2014).
21. SOURDIN, supra note 18, at 10; Sourdin et al., supra note 18.
22. SOURDIN, supra note 18, at 129.
‘improper’ influence on a judge.”

Monika Zalnieriute and Felicity Bell distinguish between two waves of AI that can be applied to judicial decision making. The “first wave” of AI is a process that follows a series of pre-programmed rules to mirror the response of a human expert. The “second wave” of AI includes techniques such as supervised machine learning and deep learning which enable systems to “learn” from data in order to draw inferences about new situations. In the context of the justice sector, machine learning can have a significant impact on litigation outcome prediction, the automated examination of legal documents, and the analysis of factual matters within a legal context. In addition, the digitization of court records provides a rich database that can be used in a machine learning environment, and it is for this reason that court-related AI developments are closely linked to the digitization of court materials. Developments in this AI area are predicted to become even more extensive in the coming years partly because of developments in the deep learning area which enable high-level data features to be analysed using artificial neural networks.

In terms of adjudicative functions, as noted, there are already some examples of AI informing human decision making in the justice sector. In China, the USA, and other jurisdictions, AI is already changing judicial decision making. In the legal sector, there are numerous predictive analytics developments that enable forecasts to be made regarding the outcome of litigation. And the rise of more sophisticated

---

26. Id.
28. Sourdin, supra note 18, at 128.
29. See KEVIN D ASHLEY, ARTIFICIAL INTELLIGENCE AND LEGAL ANALYTICS (2017) (giving a more complete description of these processes and systems).
chat and text generation systems such as GPT 3.5 and 4 will continue to foster discussion about the replacement of judges by forms of AI. The impacts of these technologies are currently emerging in some civil dispute areas and are expected to have more significant future impacts, especially in the criminal law area.\textsuperscript{32}

Notably, the process of automation is “characterised by a continuum of levels rather than as an all-or-none concept,”\textsuperscript{33} whereby decisions may be wholly or partly automated and automation may occur at different stages of a decision making process with different degrees of human oversight and verification involved.\textsuperscript{34} The extent to which systems of automation will be implemented is dependent on ‘in country’ cultural, judicial, and political factors, as well as technological readiness.\textsuperscript{35} In addition, automotive technological developments give rise to judicial independence issues related to the potential for executive control or interference. To this point, it can be anticipated that the reception of such developments by judges will be dependent on whether they were led by courts and Chief Justices (as has occurred in the USA), supported by organizations external to the court and government, or arose from a combination of judicial officers working with the government.\textsuperscript{36}

In some instances, the consideration of reforms has been considered from a broader policy perspective. In the EU, a proposed AI Act notes (in Annexure II) that certain judicial activities are automatically described as “high risk” and will therefore be subject to a different regulatory regime; these activities are described as “AI systems intended to assist a judicial authority in researching and interpreting facts and the law and...
in applying the law to a concrete set of facts.” The proposed regulatory approach to such activities is, however, somewhat unclear. In particular, there are issues about how the AI Act can adapt to a rapidly evolving technological age and, more importantly, from a social constitutional framework how the impacts of technology transcend notions of individual or group justice and impact a range of norms, behaviours, and approaches.

III. THE ROLE OF JUDGES IN CONTEMPORARY SOCIETY

Of course, the emergence of Judge AI raises questions concerning the role of judges in contemporary society. Such a role can be described as “multi-faceted, complex and subject to significant jurisdictional variation.” In addition to adjudicatory functions, the judicial role can incorporate activism, complex interactions with people, dispute settlement, case management, public and specific education activities, and social commentary, as well as adjudicatory functions, which might be conducted with other judges or, less commonly, in some jurisdictions, with lay people (juries).

The role and approach of judges when dealing with disputes can also vary significantly, with some modern trends in judicial approaches leading some judges to be more “responsive,” as discussed further below. In addition, some judges are also much more involved in “creating” the law. Judicial systems that support the ongoing development of the law through the creation of precedent, which necessarily involves both creativity and an understanding of social change, may raise very different issues relating to the application of technology and AI to the judicial role. This can be contrasted with judicial systems which may only support limited creativity or which are bound by a limited consideration of contextual factors.

Indeed, the role of a judge varies considerably across jurisdictions and is linked to the political systems and traditions within which judges

37. Sourdin, supra note 18, at 32.
38. Sourdin, supra note 18, at 10; Sourdin et al., Supra note 18, at 88; see also Tania Sourdin & Archie Ziski, The Multi-Tasking Judge: Comparative Judicial Dispute Resolution (2013) (for a helpful discussion).
40. Sourdin, supra note 18, at 34.
41. Id. at 34.
operate. In most democratic countries, the concept of separation of powers informs the role of the judiciary as one of the three arms of government—the other two being the executive and the legislature. The judicial arm is intended to support a system of checks and balances, to enforce legislative requirements, and to also support the legality and appropriateness of government decision making. It has been said that this role must not be overlooked by taking an “unreasonably narrow view of the judiciary’s role, seeming to assume that the judiciary’s only role is to robotically apply the law to the facts.” A central tenet of judicial independence that is challenged by Judge AI relates to how an AI Judge could really serve as a significant ‘check’ on the executive and legislative branches of government.

IV. THE ROLE, NATURE, AND IMPACT OF ‘DIGITAL SWITZERLANDS’ IN JUSTICE SYSTEMS

An issue linked to judicial independence that may be separate or intertwined with the potential for executive control or interference discussed above is the development, ongoing role, and nature of “digital Switzerland” in society which often operate without any clear jurisdictional boundaries. The term ‘digital Switzerland’ refers to the digital tech giants such as “Facebook, YouTube, Twitter and Co.” Commentators have voiced concerns regarding digital overreach in that many digital corporations may not understand democratic notions or the rule of law and may reshape judicial activities in unforeseen ways through their technologies. Indeed, some question how the law is currently being developed and query the extent to which it is already reflecting the interests of digital giants. As Paul Nemitz explains:

[The] explicit or implicit claim [of digital corporations, activists, programmers] that parliamentarians and governments do not understand the Internet and new technology such as AI, and thus have no legitimacy to

42. Id. at 48.
43. Id. at 189.
44. Id.
46. SOURDIN, supra note 18, at 189.
47. Paul Nemitz, Constitutional Democracy and Technology in the Age of Artificial Intelligence, 376 PHIL. TRANSACTIONS ROYAL Soc’y, Nov. 2018, at 1, 1.
48. SOURDIN, supra note 18, at 196–97.
49. See, e.g., Nemitz, supra note 47.
put rules for these in place, is not matched with a self-reflection on how little technologists actually understand democracy and the functioning of the rule of law as well as the need to protect fundamental rights in a world in which technology increasingly tends to undermine all these three pillars of constitutional democracy.\textsuperscript{50}

This issue is relevant as “digital Switzerlands” and large corporations are already reshaping what might be dealt with in domestic legal systems. For example, shifts towards private arbitration, or even international commercial arbitration, can be viewed as a shift away from State based courts to privatised justice systems. While some may be supportive of such developments, there are questions about how far such privatization could go and to what extent citizen rights should be determined via private mechanisms. Those systems that are created may not be aligned with societal or other interests, may not be independent nor engaged in reform measures that support democratic engagement, and may corrode the current foundational understanding of “how the international public represents itself through law and legal institutions, which include both states and international organizations.”\textsuperscript{51} Where private interests control data and also the means with which to create sophisticated AI systems, there is a danger that the state becomes “hollowed out”\textsuperscript{52}—whereby a core set of functions traditionally performed by the state are privatized and the “co-constitution,”\textsuperscript{53} which exists between international law and sovereign states, in which limits are set on the role outsourced actors may perform, is broken.\textsuperscript{54}

Judges, to address the issues that can be linked to executive or “digital Switzerland” overreach, themselves must be responsive and engaged in the technological challenges that will take place in courts (and beyond in the broader justice system) and undertake the important policy and reform work that relates to advanced technological design. In doing so, judges perform state functions as expected by society through exercising its “public laws adopted within national and international

\textsuperscript{50} Id. at 5.
\textsuperscript{51} Melissa Durkee, \textit{Introduction to the Symposium on Frédéric Mégret, “Are There ‘Inherently Sovereign Functions’ in International Law?”115 AMERICAN J. INT’L L. UNBOUND 299, 300 (2021)}.
\textsuperscript{52} Frédéric Mégret, \textit{Are There “Inherently Sovereign Functions” in International Law?}, 115 AMERICAN J. INT’L L. UNBOUND 452, 452 (2021).
\textsuperscript{53} Durkee, \textit{supra} note 51, at 299.
\textsuperscript{54} Mégret, \textit{supra} note 52, at 458; see Durkee, \textit{supra} note 51, at 299 (arguing that those things which are ‘inherently sovereign’ to the state are open to debate).
ROBO JUSTICE: CONSTITUTIONAL ISSUES WITH JUDGE AI

constitutional frameworks,” which can be social as well as legal. Such judicial performance also emphasizes the tenant of self-government—one so foundational to democratic societies and the social constitution under which those who are subject to the law are also its authors. As such, there is a need to appoint judges with backgrounds that include sophisticated understandings of new technologies and the time and ability to contribute to the design systems that are responsive to judicial and user needs. Should such system design and operation instead fall entirely on the executive arm of government or external consultants without adequate supervision or understanding of judicial functions and approaches, more significant issues regarding the separation of powers, judicial independence, technological overreach, and overreliance all emerge, thus eroding the social constitution in which accountability, representativeness, and self-government are enshrined.

In an expanded formulation of the judicial role, it is human judges who support and maintain important aspects of society and continue to preserve and, to some extent, protect against approaches that may oversimplify discussion and erode the rule of law as has taken place “by a group of companies which concentrate power in their hands without precedent in history.” In this regard, the inability of some judges to appropriately use technology to engage in discussion within society is problematic and, as I have previously noted, if activism (or responsiveness) is within the province of judicial work then it will also require additional consideration of how judicial views can be expressed in a modern society.

There are also concerns that ethical decision-making is in the hands of digital tech giants who may have a variety of incentives to develop technologies that do not support social good or human wellbeing. In this sense, judges could therefore be perceived as critical in ensuring that unethical and unlawful AI use is reduced. For example, replacing human judges with Judge AI could mean the gatekeepers of AI development are other forms of AI, giving rise to concerns regarding the

55. Id.
57. SOURDIN, supra note 18, at 197.
58. See id. at 189–209; Durkee, supra note 51, at 301.
60. SOURDIN, supra note 18, at 197–199.
operation of undetected biased or inappropriate forms of AI.\textsuperscript{62}

\section*{V. A CHALLENGE TO CONVENTIONAL THEORETICAL GOVERNANCE APPROACHES}

The varying and developing relationships between judges, courts, and emergent technologies challenge conventional theoretical approaches to governance—particularly the separation of powers doctrine, as discussed above. Should governments play a substantive role in setting up AI systems that replace judges (or even “support” them through the use of “template” decisions), it could be suggested that there is an in-built bias towards the government.\textsuperscript{63} These issues can be linked to concerns regarding “the transparency or opacity of algorithmic decision making.”\textsuperscript{64}

There has been some discussion relating to automated decision making by the government and some limited analysis of the role a judge plays in reviewing an automated decision, however, much of that analysis assumes that a human judge rather than a form of AI would review an automated decision. According to Monika Zalnieriute et al., automation has the potential to make government decision making more accurate, efficient, and fair through its ability to efficiently explain precisely how every variable in a decision was set and why each conclusion was reached.\textsuperscript{65} However, Cary Coglianese and David Lehr have discussed the “black box” problem of machine learning: that machine learning results that are not intuitively explainable “cannot support causal explanations of the kind that underlie the reasons traditionally offered to justify governmental action.”\textsuperscript{66} Zalnieriute et al. note that “[t]he alignment of automated government decision-making with rule of law values hinges on the appropriateness of design choices”\textsuperscript{67} whereby automated systems use explicit rules written by humans as opposed to rules derived empirically from historical data.\textsuperscript{68}


\textsuperscript{63} SOURDIN, supra note 18, at 195.

\textsuperscript{64} Id. at 72.


\textsuperscript{67} Zalnieriute et al., supra note 65, at 428.

\textsuperscript{68} Id.
Ensuring human involvement by independently justifying the decision and facilitating appeal processes is one way in which accountability can be preserved.69

A further challenge that arises relates to the legality of AI decisions. Questions such as who makes the decision in the context of automated decision making as well as who possesses the legal authority to do so arise.70 With the proliferation of machine learning algorithms, “public officials, lawyers, and scholars will confront choices about whether to encourage or constrain this technology.”71 Writing extra-crucially, Justice Melissa Perry of the Federal Court of Australia has warned that:

It cannot be assumed that a statutory authority vested in a senior public servant which extends by implication to a properly authorised officer, will also extend to an automated system; nor that authority to delegate to a human decision-maker will permit ‘delegation’ to an automated system.72

According to Zalnieriute and Bell, automation also has the potential to undermine the independence of the judiciary, even where it is intended to assist a judge.73 Judicial independence is defined as encompassing: (i) the independence of the individual judge or decisional independence; (ii) independence of the judiciary as an institution from interference or usurpation by the other branches of government; and (iii) independence afforded by administrative and fiscal self-management.74 Zalnieriute and Bell argue that judicial independence could be undermined where the automated tool that is relied upon to assist judges uses proprietary software developed by a private company that is protected by intellectual property laws, making it impossible to understand how its outputs have been generated.75

In addition, as Coglianese and Lehr have noted, citizens tend to view governmental institutions as more legitimate when they operate...

69. Id. at 445.
70. Sourdin, supra note 11, at 1126.
71. Coglianese & Lehr, supra note 34, at 1176 (discussing the use of artificial intelligence in administrative application).
72. Perry, supra note 34, at 31.
74. Sourdin, supra note 18, at 90.
75. See Zalnieriute & Bell, supra note 73.
with understanding and empathy. They argue that the idea of the government “reducing individuals to data points that are then fed into an algorithm will seem disconcertingly impersonal—even if ultimately more accurate and efficient.” Here, issues of algorithmic bias are particularly relevant as often administrative decision making can impact the most vulnerable members of society, as discussed later in this article.

In addition, apart from the issues involved in the adoption of a Judge AI approach, there are practical difficulties that have been explored by some writers and may be partly dependent on whether conditions relating to strong or weak judicial discretion apply. In one analysis of European Court of Human Rights decisions, it was noted that there are substantive issues with systems that are able to do more than classify rather than forecast judicial decisions.

VI. DEVELOPING A CONTEMPORARY APPROACH TO JUSTICE AND AI JUDGING

A focus on social interaction enables exploration of how adaptation through judicial responsiveness may support the development of ethical approaches that respond to vulnerable populations and social needs. Such an approach can also be used to inform constitutional reform by focusing on the role of judges from a “responsive” rather than an “activist” lens.

VII. JUDICIAL RESPONSIVENESS

Judges, as both myself and Archie Zariski have noted, have always been responsive. That is, they: (i) finalize all disputes before them; (ii) consider the submissions of litigants; (iii) use submissions in making decisions; and (iv) explain and justify their decisions. The classic or passive responsiveness which previously dominated many non-inquisitorial Western legal systems can be contrasted with a more

76. Coglianese & Lehr, supra note 34, at 1219 (The use of artificial intelligence in administrative application); see also Sofia Ranchordas, Empathy in the Digital Administrative State, DUKE L. REV. (forthcoming May 2022).
77. Coglianese & Lehr, supra note 34, at 1219 (discussing the use of artificial intelligence in administrative application).
78. Masha Medvedeva et al., Automatic Judgement Forecasting for Pending Applications of the European Court of Human Rights, in PROCEEDINGS OF THE FIFTH WORKSHOP ON AUTOMATEC SEMANTIC ANALYSIS OF INFORMATION IN LEGAL TEXT ASAIL 2021 CEUR WORKSHOP PROCEEDINGS 12–23 (K. D. Ashley et al., eds., 2021).
79. Sourdin et al., Supra note 18.
80. Id. at 1.
progressive or active model of judging. This latter model, further to the above elements, includes at least one of the following: (i) responsiveness to accountability for public investment in the legal system and the demand for justice from it; (ii) responsiveness to the problems of an interdependent, network society; (iii) responsiveness to litigants' experiences of the legal system and courts; and (iv) interactive responsiveness in the context of public attention.\(^\text{81}\) Responsiveness, it can be said, is grounded in a sense of empathy which ultimately enhances societal trust in the law and makes public administration “more humane, democratic, and legitimate.”\(^\text{82}\) After all, “[e]mpathy is an important component of social cognition that contributes to one’s ability to understand and respond adaptively to others’ emotions, succeed in emotional communication, and promote prosocial behavior”\(^\text{83}\) (emphasis added).

Further, apart from “adhering to the essential principles of fidelity to law, impartiality, and integrity,”\(^\text{84}\) today’s “fully responsive judge” is also a cost-conscious manager of litigation with a concern for ensuring access to justice; a quick learner with the curiosity and patience to inquire into the foreseeable consequences of their decisions; a student of human nature who values and works at establishing respectful relations with litigants and colleagues; and a public figure comfortable in the roles of ambassador for justice and public legal educator.\(^\text{85}\) Ronald Sackville, former Justice of the Supreme Court of New South Wales in Australia has commented on the idea of judicial responsiveness, noting that this concept:

\[\text{[E]ncapsulates the idea that the courts should accept responsibility not merely for managing the conduct of litigation, but for a wider range of activities designed to enhance the responsiveness and accountability of the legal system to the community, but in ways that are consistent with judicial independence.}\(^\text{86}\)

To support technological change, a clearer appreciation of the

---

81. Id. at 2.
82. See also Ranchordas, supra note 76.
84. SOURDIN, Supra note 18, at 47.
85. Sourdin et al., Supra note 18.
judicial function is required, as well as a recognition that the function is not only multi-faceted but also responsive in the broadest definitional sense. This approach informs a more effective analysis of the role of AI and judging, particularly when considering the extent to which judges can adapt to the introduction of more disruptive technologies.

Morison and Harkens have similarly noted that most AI researchers fail to recognize what courts and judges actually do.87 Specifically, they note that judging “is not a singular activity with a fundamental method that is unchanging across whatever context it is being employed in.”88 Significantly, Morison and Harkens conclude that the ‘social aspect’ of the judicial role means that while new technologies may disrupt current working patterns, they cannot produce a new kind of justice system alone.89 For a new type of justice system to develop, much more will be required. Changing judicial roles and functions will not, in isolation from other changes to courts and the attendant cultures of the legal profession, enhance access to or support a more effective justice system. In addition, judicial cultures that are varied and often idiosyncratic will need to be considered in the context of which ‘human’ functions should be retained in the context of technological developments, particularly in the AI area.90 Despite some judges eschewing such human traits as empathy and social skills, these may support the retention of human judges, as technology is currently very poor at replicating them.91 Such behaviors focus on the ability of human judges to more effectively enable litigants to determine outcomes through the judge demonstrating “curiosity, emotional understanding of parties and their lawyers, and their agile questioning and exploration of issues.”92 Certainly, these behaviors are inherent in the broader activities involved in responsive judging.

The role of judges as social commentators must also be recognized and grappled with when considering how AI can replace human judges. Morison and Harkens, for example, have noted that this role incorporates resistance and contestation for “essentially human values” in an area that “maintains a central social element.”93

88. Id. at 629.
89. Id. at 631.
90. SOURDIN, supra note 18, at 176–78.
91. See Ranchordas, supra note 82.
92. SOURDIN, supra note 18, at 10 (2021); Sourdin et al., Supra note 18, at 97.
By reframing the definition of “justice” to include an overarching focus on promoting human wellbeing, a restrained and thoughtful approach to the use of more disruptive technologies in the justice sector can take place. This is essential because without clear justice frameworks that incorporate value-related material, such as those linked to requirements of human dignity and wellbeing, AI developments are likely to result in arrangements that do not support justice objectives.\textsuperscript{94}

“Wellbeing” is increasingly noted as relevant in terms of the objectives and purpose of the justice system, as well as ethical frameworks in respect of AI. For example, in May 2019 the OECD Recommendation of the Council on Artificial Intelligence recommended the adoption of values-based principles “for the responsible stewardship of trustworthy AI.”\textsuperscript{95} Its issued statement on the purposes of AI included wellbeing.\textsuperscript{96} There is an emerging literature relating to the meaning of well-being and, more specifically, in the context of government policy.\textsuperscript{97} The literature suggests that although wellbeing is a subjective matter, it includes positive attitudinal factors,\textsuperscript{98} life and health satisfaction,\textsuperscript{99} confidence and a sense of fairness,\textsuperscript{100} and may also refer to sustainability.\textsuperscript{101} A number of writers have noted that wellbeing is a multidimensional concept and the use of the term by many governments is indicative of a changed way of thinking in that “[T]he wellbeing of citizens, not the wellbeing of their bank accounts, was considered to be

\textsuperscript{94} Berkeley Dietvorst, Joseph Simmons & Cade Massey, Overcoming Algorithm Aversion: People will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them, MAN. SCI., 115, 1155 (2016) (reduction in algorithm aversion by allowing an individual's control over an imperfect algorithm's forecast).

\textsuperscript{95} Organisation for Economic Co-operation and Development, Recommendation of the Council on Artificial Intelligence (OECD Legal Instruments 0449, May 2019).

\textsuperscript{96} Rachel Dodge et al., The Challenge of Defining Wellbeing, 2 INT'L J.W. 222, 222 (2012).


\textsuperscript{100} Kong Weng Ho, Which Class and What Squeezes? Relationships with Wellbeing, National Pride, and Inequality (Working Paper, July 16, 2015).

\textsuperscript{101} Necati Aydin, Subjective Well-Being and Sustainable Consumption, 6 INT'L J.E.C.E.S.S. 133, 133 (2010).
the end goal of government.”  

For example, as the Australian Productivity Commission noted in its 2014 Report on Access to Justice Arrangements, policy decisions about justice require consideration of the core overriding objective of the justice system, which is “to enhance community wellbeing.”  

In this sense, economic and financial wellbeing are perceived to be a subset of wellbeing.

There are also different philosophical understandings surrounding the meaning of “justice.”  

The traditional view, likely to be adopted by litigation supporters, is that justice can only take place within the courts, where a judge is able to articulate understandings about the rule of law.  

The alternate view which I adopt is a broader conception of justice that extends from the role the judiciary plays in the public adjudication of civil disputes to “the relationships . . . between people and their ethical values.”  

As a result of the varying emphasis in different countries on well-being and different understandings of justice, it is perhaps not surprising that in some AI ethical frameworks such concepts are more strongly emphasized.

Other conceptions of justice include references to procedural, substantive, distributive, transparent, enforceable, and sustainable components, each of which can be considered from an individual, group, and societal perspective. While some of these conceptions can be narrowly construed, the understanding and meaning of such concepts is critical in considering how technological change may impact the structures, systems, and processes within the justice system. Indeed, in this regard, they provide a framework for deeper analysis and evaluation.

Of course, in reframing the definition of justice, one must investigate what justice determined by judges or “juridical justice” as it

---

106. Sourdin et al, supra note 13, at 17.
107. Id. at 21–22.
has been termed actually is—separate from justice as a political or moral project. Teubner suggests that justice is “no longer only a construct in philosophical discourse, but is reconfigured in concrete social practices, such as in litigation, contracting, standard-setting, and legislation, in the incessantly changing self-images of the practice of law.” Therefore, juridical justice is critical in law’s “internal self-reference,” which applies past decisions to new factual situations but also evolves through the external social processes that redirect the juridical semantics of justice.

In this issue, Giovanni De Gregorio notes that technological changes may create a new system of digital norms that can “escape the logic of the rule of law” and in doing so can threaten substantive justice by reshaping the norms that are accepted as part of the juridical justice. This approach which considers technological change through a rule of law perspective ensures that in rethinking frameworks in this area, a perspective beyond individual and group justice is enabled with a focus on how normative arrangements may not be fostered by a community but led by programmers, developers, and service providers. Other commentators have raised this issue by considering how boundaries can be maintained so that important features relating to legal reasoning are supported.

For the potential negative consequences of Judge AI to be mitigated for the delivery of juridical justice, courts and judges can develop a responsiveness that focuses on what justice they deliver and how it is delivered, and critically examine the current restrictiveness of some judicial approaches. There may be self-limitations that require resolution to enable an “expansionist course of action” to operate which broadens the social constitution to embrace the positive potential of Judge AI to legal practice. Juridical justice, separate from moral or political justice, has the double requirement of both internal consistency as well as responsiveness to the demands of the social environment in which it exists.

111. Id. at 3.
112. Id. at 7.
113. Christopher Markou & Simon Deakin, Ex Machina Lex: Exploring the Limits of Legal Computability, in IS LAW COMPUTABLE? CRITICAL PERSPECTIVES ON LAW + ARTIFICIAL INTELLIGENCE (Simon Deakin & Christopher Markou eds., 2020).
114. Id. at 6.
Ethical frameworks which provide guidance throughout the lifecycle of forms of AI can support more sophisticated and appropriate justice mechanisms. In general, ethical concerns about AI and technology use can be grouped into five broad categories: (i) fairness, transparency, and explainability; (ii) responsibility and accountability; (iii) robustness and reliability; (iv) privacy and trust; and (v) safety and security.116

There are different points at which ethical decisions regarding technology and AI use arise: in the design and development stage as well as the application and implementation stage, ethical decisions may need to be made.117 In addition, ethical concerns can be contextual in nature with the use of technology may be ethically appropriate when used by a judge in one context but not another.118 In addition, jurisdictional differences mean that what may be considered as ethical or an appropriate use of technology in one jurisdiction may not be considered that way in another.119

“The underlying values and objectives of the justice system in each jurisdiction can help in determining whether an action is ethical or not, and decisions about how to construct and interpret national and regional ethical guidelines can also be relevant.”120 This values-based approach must be incorporated into an ethical framework that reflects the “values underpinning the judicial system within a particular jurisdiction.”121

In many jurisdictions, values relating to human rights will be dominating features of an ethical framework. By way of example, in December 2018 the Council of Europe European Commission for the Efficiency of Justice developed perhaps one of the most relevant documents in relation to Judge AI: the “European Ethical Charter on the use of Artificial Intelligence in judicial systems and their

117. Sourdin, supra note 18, at 257.
118. Id. at 236.
119. The author notes that what might be considered to be palatable by the courts in China in the context of social surveillance data and privacy, might not be considered in the same way in another jurisdiction.
120. Sourdin, supra note 18, at 236 (2021).
121. Id. at 257; see also Matthew Beard & Simon Longstaff, Ethical by Design: Principles for Good Technology, The Ethics Centre Australia (Sep. 2018), https://ethics.org.au/ethical-by-design/.
environment.”122 The five key principles adopted by the CEPEJ are stated as follows

1. Principle of respect for fundamental rights: ensure that the design and implementation of artificial intelligence tools and services are compatible with fundamental rights.

2. Principle of non-discrimination: specifically prevent the development or intensification of any discrimination between individuals or groups of individuals.

3. Principle of quality and security: with regard to the processing of judicial decisions and data, use certified sources and intangible data with models elaborated in a multi-disciplinary manner, in a secure technological environment.


5. Principle ‘under user control’: preclude a prescriptive approach and ensure that users are informed actors and in control of the choices made.123

In other jurisdictions, other values may be emphasized differently. Of course, the rule of law is not a static concept. It evolves in response to changing societal values and the operation of government. As technology reshapes society and government interacts with the community, it can be expected in turn that our understanding of the rule of law will shift. Values such as transparency and accountability, predictability, consistency, and equality before the law may remain central to conceptions of the rule of law and social constitution, but their interpretation and application may change.124 The benefits offered by such technologies, such as their capacity to reduce government

123. Id.
spending, may be so significant as to demand greater accommodation within the rule of law framework.\textsuperscript{125}

It has been argued that “human dignity” should be a guiding principle when considering how technology can be utilized by judges.\textsuperscript{126} According to Kaminski, “allowing a decision about humans to be made by a machine inherently treats humans as objects, showing deep, inherent disrespect for peoples’ humanity.”\textsuperscript{127} Further, Kaminski argues that if algorithmic decision making “does not allow individuals to proclaim their individuality . . . then it violates their dignity and objectifies individuals as their traits, rather than treating an individual as a whole person.”\textsuperscript{128} Butler has observed that judging is “a very human endeavour, reflecting all the variation in experience, perspective, humanity, common sense, and understanding of the law of the judges themselves.”\textsuperscript{129} In addition, citizens tend to view governmental institutions as more legitimate when they operate with understanding and empathy.\textsuperscript{130}

However, Volokh has also noted that although an AI Judge may not be capable of compassion or mercy, the focus needs to be on judgments that possess these qualities rather than decision makers that do so.\textsuperscript{131} If there is a focus on AI judgments (rather than the AI Judge) it is possible that, at some point in the future, an AI judgment could demonstrate compassion, mercy, and wisdom, and perhaps even more effectively than a human judge might.\textsuperscript{132} Further, AI systems could be developed, acknowledging the value (and cost) that human judges can “add” in retaining humanity in decision making and preserving the dignity of humans.\textsuperscript{133}

The author has developed ethical principles adapted from the above-mentioned concerns about AI and technology use which could be built into an ethical framework cognizant of the decisions to be made regarding AI use over its lifecycle:

\textsuperscript{125} Id.
\textsuperscript{127} Id. at 1542.
\textsuperscript{128} Id.
\textsuperscript{129} Id.
\textsuperscript{131} Coglianese & Lehr, \textit{supra} note 34, at 1219 (The use of artificial intelligence in administrative application); Ranchordas, \textit{supra} note 76.
\textsuperscript{133} Id.
• Human, social, environmental well-being and sustainability: Throughout their lifecycle, Judge AI systems should benefit individuals, society and the environment. They should promote well-being, support justice system objectives and the sustainability of the justice system.

• Human-centered values: Throughout their lifecycle, Judge AI systems should respect human rights, diversity and the dignity and autonomy of individuals. They should also promote trust in, and the independence of, the judicial system.

• Fairness: Throughout their lifecycle, Judge AI systems should be inclusive and accessible, and should not involve or result in unfair discrimination against individuals, communities or groups.

• Privacy protection and security: Throughout their lifecycle, Judge AI systems should respect and uphold privacy rights and data protection, and ensure the security of data.

• Reliability and safety: Throughout their lifecycle, Judge AI systems should reliably operate in accordance with their intended purpose.

• Transparency and explainability: There should be transparency and responsible disclosure to ensure people know when they are being significantly impacted by a Judge AI system, and so that they can find out how to engage with a Judge AI system. Decisions produced as a result of Judge AI should explain not only how the decision was reached in terms of the evidence and the law but also how the AI system developed the decision.

• Contestability: When an AI system significantly impacts a person, community, group or environment, there should be a timely process to allow people to challenge the use or output of the Judge AI system. This requires that simple review processes be developed so that a Judge AI decision is not binding if any person objects to this within a set timeframe and so that referral to a
human judge can take place.

- **Accountability:** Those responsible for the different phases of the Judge AI system lifecycle should be identifiable and accountable for the outcomes of the Judge AI system. Human judicial oversight of the systems should be enabled and supported.\(^{134}\)

X. **CHALLENGES PRESENTED BY THE INCREASED USE OF JUDGE AI AND TECHNOLOGY IN THE JUSTICE SYSTEM**

The challenges presented by increased technological developments in a global context are readily apparent in the context of differing justice conceptions and a focus on “fast,” “low cost” justice delivery, perhaps in the absence of justice itself. Margaret Beazley, former President of the New South Wales Court of Appeal in Australia, has argued that we must ensure “that technological change supports the administration of justice and the rule of law and does not supplant it with concepts that are alien to our rich legal tradition.”\(^{135}\) Similarly, Justice See Kee Oon of the Supreme Court of Singapore has argued that “while technology can lead to greater efficiency and enhance the delivery of justice, the human touch remains essential. In the delivery of justice, human experience, empathy, and common sense reasoning play a critical role.”\(^{136}\)

The benefits of technological development in law and the creation of an e-justice system are evident. A cohesive e-justice system involves online courts as well as triaging, ADR and ODR, and engagement with justice apps that may be of assistance.\(^{137}\) This system is aimed at “improving service delivery and collaboration between all justice players

---


\(^{135}\) Margaret Beazley, President, New South Wales Court of Appeal, Speech at the State of the Profession Address, New South Wales Young Lawyers Sydney, Law in the Age of the Algorithm (Sept. 21, 2017).

\(^{136}\) Justice See Kee Oon, Speech at the State Courts Workplan 2019, State Courts: 2020 and Beyond (Mar. 8, 2019).

through the use of [information and communication technologies].”

A 2016 panel discussion held by the United Nations and the Rule of Law Unit specifically highlighted the benefits of e-justice: financial and time savings through efficient and effective justice delivery, improved accessibility of information, greater transparency of the judicial system, and increased quality of justice brought about as a result. The panel also found that the ability to monitor proceedings, good quality legal databases, and availability of public versions of judgments provided an opportunity for public awareness and scrutiny which enhanced the accountability of the judicial system.

However, there are a number of challenges I have identified which are associated with the use of e-justice. These include the difficulty in ensuring the validation and authentication of information as well as data security. In addition, a UK Justice Committee report on the shift to online hearings in the UK as a result of the COVID-19 pandemic has confirmed that online hearing approaches are not suitable for all people and that the move to online communications can exacerbate the disadvantages faced by “vulnerable users.” It has been recognized that “poor digital skills, limited access to technology, low levels of literacy and personal disadvantage experienced by particular groups” create barriers to accessing digital justice services, adequate legal advice, and support.

Robert Condlin has also questioned whether “the cheap and efficient processing of disputes is a capitulation to the conditions of modern society more than a superior system for administering justice.” In this regard, it has also been suggested that more extensive ODR systems may restrict the ability of parties to argue the substantive merits of their claims, whereby uncoupling disputes from their substantive merits


140. Id.

141. See SOURDIN, supra note 18, at 173–74.

142. E-justice: Enhancing Transparency, Effectiveness and Access to Justice, supra note 139.


144. Id.

undermines the fairness of individual outcomes and, if widespread, threaten the legitimacy of dispute resolution systems themselves.\textsuperscript{146}

Challenges surrounding the existence of a digital divide are also central to any discussion about the potential for new technologies to improve access to justice. As internet use continually grows, it is often assumed that people will be able to access ODR options via the internet or use supportive or other technologies.\textsuperscript{147} The existence, however, of a “digital divide” means this is not always the case.\textsuperscript{148} There is evidence that the digital divide in internet connectivity is shrinking as access to technology increases. Indeed, in 2020, mobile phones became the dominant means by which people access the internet around the world, with an estimated 62.6% of the world’s population, or 4.78 billion people, using a mobile phone in 2020.\textsuperscript{149}

However, Li, Burke, and I have similarly identified an “uneven” readiness to adopt new technologies that can be associated with geographical location, age, economic circumstances, and other factors that can be linked to vulnerability.\textsuperscript{150} Any meaningful implementation of justice technologies, therefore, requires consideration of the socioeconomic and other barriers to technology access.\textsuperscript{151}

Digital literacy issues also pose a challenge when considering access to justice that is driven by newer technologies.\textsuperscript{152} A 2008 report prepared for the Law Foundation of Ontario found that “internet and other text-based solutions are of limited use to people who do not have the literacy skills to use them or to use them effectively.”\textsuperscript{153} The report cautioned that “vulnerable people, because they face language barriers, isolation, poverty, or a cluster of other difficulties that often accompany a legal problem, [ideally] need to receive direct services rather than to rely on self-help [through either digital or paper-based resources].”\textsuperscript{154}

\begin{itemize}
  \item \textsuperscript{146} Id. at 722; see generally Hazel Genn, \textit{What Is Civil Justice For? Reform, ADR, and Access to Justice}, 24 (2012) YALE J. L. & HUMANS. 397, 397-417 (discussing civil justice policy).
  \item \textsuperscript{147} SOURDIN, supra note 18, at 29.
  \item \textsuperscript{150} SOURDIN et al, supra note 13, at 34.
  \item \textsuperscript{151} SOURDIN, supra note 18, at 182.
  \item \textsuperscript{152} Id.; Jeff Giddings & Michael Robertson, \textit{Informed Litigants with Nowhere to Go: Self-Help Legal Aid Services in Australia}, 26 ALTERNATE L. REV. 184, 188 (2001).
  \item \textsuperscript{153} KAREN COHL & GEORGE THOMSON, \textit{Connecting Across Language and Distance: Linguistic and Rural Access to Legal Information and Services} 35 (2008).
  \item \textsuperscript{154} Id. at 52.
\end{itemize}
This was confirmed by the findings of the UK Justice Committee report discussed above.

XI. PROPOSALS FOR THE CREATION OF A CONSTITUTIONAL FRAMEWORK FOR FACILITATING JUSTICE IN CYBERSPACE

Exploring how judges can be involved in the design of a constitutional framework to facilitate justice in cyberspace is critical in ensuring that constitutional approaches respond to the demands raised by technologies that already exist as well as those developments that will reshape society in the years to come. This involves considering developments that will occur within as well as outside courts and relate to the judicial role and function as discussed above. In addition, active judicial engagement and adaptability are considered in the context of how reform can take place and what ethical guidance is required.

XI. JUDGE-LED REFORM THROUGH AN ETHICAL FRAMEWORK

Online courts may be developed without the involvement of judges at all—which raises issues regarding what a court is and what role, if any, a judge must play in order for an entity to be called a “court.”¹⁵⁵ One type of online court which may be developed is through existing judges managing and developing the courts together with court processes-related apps and bots. Recent discourse suggests that the impact of COVID-19 has caused a significant increase in mobile app downloads and usage.¹⁵⁶ Such results yield the inevitable conclusion that courts, to remain relevant, must consider how apps can be developed that enable people to engage with them. In many places, there is a thriving new industry that is focused on justice app and bot development which is oriented towards the provision of cheaper and faster legal advice; the triaging of disputes to lawyers, mediators, and others; the finalization of disputes via ODR; the preparation and filing of court documents; the furnishing of advice relating to potential courts outcomes; the interpretation of financial records; and even the setting up of day-to-day arrangements relating to how and where children have contact with parents.¹⁵⁷

¹⁵⁵. See generally SOURDIN, Supra note 18.
Such reform could be led by the judiciary at a centralized level with the potential benefit of greater jurisdictional consistency.\(^\text{158}\) To some extent, this is what is already occurring in China’s “smart court” system,\(^\text{159}\) although it could be argued that it is in fact the executive arm of government embedded within the court infrastructure that is leading such developments. However, in other courts, some developments suggest that the judiciary may take a more substantive role in the creation of an online court.\(^\text{160}\) This approach requires extensive judicial input and tech savvy judges will need to be engaged in the development of the system and the reforms that are undertaken.\(^\text{161}\)

This type of reform requires input by judges and close attention to newer ways of design that explore what is useful for both judges and litigants in terms of defining and exploring a legal issue. A transition to online courts may take considerable work in terms of piloting, trialing, and developing new approaches to not simply replicate the existing paper-based approach of decades past.\(^\text{162}\) Even at the simplest level, the benefits that may flow from the use of enhanced technologies will be very limited if no adoption of a plain language approach to reform that might enable people to complete an online form or understand what is required once court proceedings are commenced.\(^\text{163}\)

A transition to online courts of course raises digital divide issues. While much of the population is online, this does not necessarily equate to digital literacy, and the evaluation of some COVID-19 arrangements has shown that there is a proportion of the population that will require additional human support to commence proceedings, defend themselves,
and continue with proceedings,\textsuperscript{164} as discussed above.

In addition, the courts of the future that will rely on newer technologies raise specific privacy and security concerns that differ from arrangements that apply to the paper-based courts of the past. There are also issues relating to how courts might consider other data such as social surveillance material and social credit rankings.\textsuperscript{165} The developments discussed above suggest that both judges and courts will need to consider their approach to reform and technology use by using ethical decision-making frameworks. Issues may arise at an individual judicial level in response to the issues of a particular case, as well as at a collective judicial level when considering overarching responses to reform. As a result, enhanced judicial education in respect of technological approaches as well as clearer ethical decision-making frameworks are necessary in the context of issues that may emerge in both online judging and Judge AI.\textsuperscript{166}

\textbf{XII. Conclusion}

There are a number of persuasive reasons why the move towards Judge AI would be a welcome one for the future justice system. First and foremost, Judge AI has the potential to increase access to justice.\textsuperscript{167} Further, it also has the potential to provide a higher quality of justice,\textsuperscript{168} and enhance the rule of law.\textsuperscript{169} Automation could make decision-making more accurate, efficient, and fair and has the potential to improve transparency and accountability in decision-making.\textsuperscript{170}

However, the future justice system requires more than an acceptance of technological innovation. It requires leadership, and if judges fail to be proactive, there is a risk that another arm of government may step in and further narrow the areas where judges

\begin{footnotesize}
\begin{enumerate}
\item \textsuperscript{164} See e.g., Natalie Byrom et al., Civil Justice Council: The Impact of COVID-19 Measures on the Civil Justice System 10 (2020); Justice Committee, Coronavirus (COVID-19): The Impact on Courts, H.O.C Doc. No.519 (Sess. 2019-21).
\item \textsuperscript{165} Sourdin, supra note 18, at 280.
\item \textsuperscript{166} Id.
\item \textsuperscript{168} Warren, supra note 167, at 227.
\item \textsuperscript{169} See, e.g., Chief Justice Helen Murrell, Speech at the Supreme Court of the Australian Capital Territory, Turn and Face the Change – New Technology and the Internationalised Judiciary, 2–3 (Feb. 1, 2016).
\item \textsuperscript{170} Zalnieriute et al., supra note 65, at 425.
\end{enumerate}
\end{footnotesize}
might operate or that public respect for the judiciary may be reduced as judges are considered to be “less relevant.”171 As such, society and judges must reconsider the role of the judge and have strategies in place to deal with the ethical and other issues raised by Judge AI.172

“Human-centered legal design” is an approach that can be used to reform the justice system, guided by ethical principles as discussed above.173 This approach could then further inform constitutional reform measures. The redesign of the justice system in terms of human-centered approaches requires judges at every stage of reform to consider litigant perspectives. It also requires courts and judges to be adaptable and to evaluate changes from a litigant’s perspective so that the justice system is “fit for purpose.”174 In addition to this, however, a “level of abstraction approach” is required for a litigant perspective in itself is not sufficient, as the role and function of a judge have implications that extend beyond individual litigants.175 Therefore, reforms must be developed and evaluated from an individual, interpersonal, group, institutional, sectoral, and societal perspective, which ensures that the various perspectives discussed throughout this article can be weighed and adequately considered.176

In this regard, the development of national and jurisdiction-specific guidelines about AI will continue to be relevant. Newer ethical guidelines that promote the notion that AI should be sustainable and note that sustainability may invoke references to governance and trust building,177 together with guidelines that consider ethical and justice design principles from a range of perspectives, will help support a “just” system.178 There is a need to develop more relatable and judge-specific ethical material about judges, technology, and AI which can assist judges in responding to issues that arise in particular case

171. SOURDIN, supra note 18, at 289.
172. Sourdin, supra note 11, at 1115–16 (2018); see also, Shang Li, Hongli Zhang, Lin Ye, Xiading Guo and Binxing Fang, MANN: A Multichannel Attentive Neural Network for Legal Judgment Prediction, 7 IEEE Access, 151144 (2019).
173. SOURDIN, supra note 18, at 289.
174. Id. at 290.
175. See id. at 236–72 (for a discussion on ‘levels of abstraction’).
176. Id. at 291.
circumstances. This in turn can assist to more accurately define judicial power in a constitutional sense.

The ongoing involvement of judges in both courts and future justice design is important as the maintenance of an independent, well-functioning judiciary is critical to the social constitution of countries in which the rule of law is central. Human judges can clearly play a role in shaping and interpreting political decision-making, correcting government actions, and preventing the abuse of power. There are significant issues with Judge AI ever being able to perform such functions, particularly where it is developed in countries where there may be little regard for the importance of the independence of the third arm of government.

These developments mean that judges must not only acquire foundational knowledge and understanding about AI, but they must also consider the implications of its use on both the justice system and the judiciary. As such, judges must have strategies in place to deal with the ethical and other issues raised by Judge AI. In particular, they must reconsider their role and to what extent it incorporates broader activism in the design of the justice system of the future, while remembering, in terms of technology and justice system reform, that just because we can do something, it does not always mean that we should.

179. Sourdin, supra note 18, at 294.
180. Id.
181. Id.
182. Id. at 295.
184. Sourdin, supra note 18, at 295.
The Transformative Potential of Meta’s Oversight Board: Strategic Litigation Within the Digital Constitution?

ANGELO JR GOLIA

ABSTRACT

Meta’s Oversight Board (OB) is at the center of divisive debates. Some commentators look at this experiment as the “Supreme Court” of a global order that is going through a process of constitutionalization. Others express concern about the OB, seen as a way to avoid public accountability and, more generally, as the legitimization of global private censorship. Either way, the debates normally focus on the important but still relatively narrow issue of content moderation and freedom of expression. In contrast, it is the manipulation of individuals and the related social effects deriving so-called informational capitalism —among them, the mental health distress of young people—that needs to be counteracted. Against this background, this article, resorting to societal constitutionalism as an analytical framework, aims to contribute to the debates on digital constitutionalism. It has two goals. First, to use the instruments available within Meta’s normative system to thematize the broader, systemic effects of social media and digitality in constitutional terms. Second, to question informational capitalism “from within,” using the right to the mental health of children as a case study. In this sense, it is an exercise of (strategic) legal imagination that focuses on the internal side of an involved actor. After the introduction, section II analyzes the features of Meta’s normative system, distinguishing between juridification (II.A) and constitutionalization (II.B). Against this background, section III outlines a litigation strategy aimed at bringing such issues before the OB. It separately examines the strategies preceding the proceeding (III.A), the authority and scope of OB’s jurisdiction (III.B), the relevant standards of review (III.C), and the potential content of the decision and the policy advisory statements (III.D). Section IV concludes.
I. Introduction

Two events marked the year 2021 for Meta, the social media company which owns Facebook and Instagram. The first event occurred on January 28, 2021. Meta’s Oversight Board (OB), the independent adjudicative body established to make consequential, precedent-setting, content moderation decisions on Facebook and Instagram, issued its first five decisions, with four out of the five overturning Meta’s actions. These decisions immediately set the tone of a sui generis creature that has since proved to be anything but docile towards its creator. The OB has been defined as “a historic endeavor both in scope and scale” and—not without some exaggeration—as “one of the most ambitious constitution-making projects of the modern era.” Be as it may, since its establishment the OB has been at the center of heated debates involving policymakers, scholars, activists, and public opinion alike. Some commentators look at this experiment as a “Supreme Court” of a global private order going through a process of constitutionalization. Others see it as a way for Meta to avoid public accountability and, more generally, as the legitimation of a worrisome form of global private censorship.
The second event was the Facebook Files scandal. Starting from mid-September 2021, the Wall Street Journal published articles on Meta based on internal documents leaked by the whistleblower Frances Haugen. Revelations included special allowances on posts from high-profile users (“XCheck”), subdued responses to flagged information on human traffickers and drug cartels, an initiative to increase pro-Meta news within user news feeds, and—most importantly for our purposes—internal knowledge on how Instagram exacerbated negative self-image in teenage girls. The documents showed that Meta had conducted internal research of how Instagram affects young users for the previous three years and was fully aware of its negative impact. The findings—confirming a growing scientific consensus—pointed to Instagram being harmful, especially to teenage girls.

---


Against the background of these apparently disconnected events,\textsuperscript{11} this article resorts to societal constitutionalism as a theory “according to which constitutional processes do not arise only within the legal systems of political-state orders, but may also emerge in private and/or hybrid legal systems of non-state social orders [and] analyzes the conditions for the emergence, co-existence, and further evolution of such constitutional processes.”\textsuperscript{12} Societal constitutionalism also aims to expand the analytical and normative scope of constitutional theory beyond political power in the narrow sense,\textsuperscript{13} to include “subtler, non-personified social processes . . . which lead to serious, widespread violations of fundamental rights.”\textsuperscript{14}

Indeed, the OB is a good candidate to apply the analytical tools of societal constitutionalism, especially when it comes to transnational corporations (TNCs).\textsuperscript{15} Likewise, the negative effects of social media on fundamental rights—especially on the right to mental health—fall within societal constitutionalism’s scope of analysis. However, it is not immediately clear how this triangulation among societal constitutionalism, the OB, and the societal effects of social media may occur. The jurisdiction of the OB is relatively limited compared to the broader impact of social media platforms. The OB is supposed to deal

---

\textsuperscript{11} On 11 October 2021, the OB announced the intention to hear Frances Haugen. Oversight Board, \textit{Oversight Board to Meet with Frances Haugen} (Oct. 2021), https://oversightboard.com/news/1232363373906301-oversight-board-to-meet-with-frances-haugen/.


\textsuperscript{14} Golia & Teubner, \textit{supra} note 13, at 376–77.

“only” with content moderation, freedom of expression, and related human rights violations. It has been designed so that broader social dynamics—especially those that cannot easily be traced back to single violations of specific rights such as explicitly racist or suicidal content—largely escape its radar.

Against this background, this article aims to explore the transformative potential of societal constitutionalism and to contribute to the broader debate on so-called digital constitutionalism. It has two goals. First, using the instruments available within Meta’s normative system to thematize the broader, systemic effects of social media and digitality in constitutional terms. Second, questioning informational capitalism from within, using the right to the mental health of young people as a case study. In this sense, it is an exercise of (strategic) legal imagination that focuses on the internal side of an involved private actor. To be sure, this is not sufficient. Limiting the unrestrained reproduction of social dynamics freed by digital technologies necessitates massive regulatory interventions by politically legitimated actors and, more generally, pressures by social actors. Further, from the perspective of the inner limits of the OB’s institutional design, ex post review can hardly address systemic issues adequately. However, tackling informational capitalism also requires strengthening the internal—both cognitive and normative—openness of the targeted social processes and, more specifically, the organizations operating within

16. OVERSIGHT BOARD, OVERSIGHT BOARD BYLAWS 5, 18, 19 (2022).


18. Broadly understood as an economic system centered around the capture and commodification of personal data for the core purpose of profit-making.” AMIRE MARE, DIGITAL SPACES, RIGHTS AND RESPONSIBILITIES 11 n. 34 (2023); see also SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER (2019) (explaining informational capitalism as a system that monetizes data). For the purpose of this article, “surveillance capitalism” and “information capitalism” are used interchangeably. However, for different approaches, especially as regards the elements of novelty and/or continuity with Fordist and neoliberal capitalism, see Amy Kapczynski, The Law of Informational Capitalism, 129 YALE L. J. 1460, 1460 (2020) (“construct[ing] an account of the ‘law of informational capitalism’”). See also Fleur Johns, "Surveillance Capitalism" and the Angst of the Petit Sovereign, 71 BRIT. J. SOCIO. 1049 (2020) (reviewing Zuboff’s The Age of Surveillance Capitalism).


them.\footnote{Cf. Teubner & Golia, supra note 13, at 15–16 (“constitutionalization cannot be based exclusively on politically legitimated norms, even when they derive from authentically deliberative processes. In the end, societal constitutionalism asks for a strategic interaction of qualitatively different kinds of norms, as affected by the digital code”).}

After this introduction, section 2 analyzes Meta’s emerging normative system, distinguishing between juridification (2.1) and constitutionalization (2.2) to assess the actual extent of such processes and offer guidance to strategic actors.\footnote{See Golia & Teubner, supra note 13, at 388–95 (discussing the normative limb of societal constitutionalism).} In this direction, section 3 outlines a litigation strategy to bring such issues before the OB, within the limits of existing rules of Meta’s normative system. It separately examines the strategies preceding the proceeding (3.1.), the authority and scope of the OB’s jurisdiction (3.2.), the relevant standards of review (3.3.), and the potential content of the decision and the policy advisory statements (3.4.). Section 4 concludes.

\section*{II. Meta’s Emerging Constitutional Elements}

trajectory of the OB closely follows the evolutionary patterns described by Teubner in his works on juridification\textsuperscript{24} and constitutionalization\textsuperscript{25} of global non-state normative systems.

\section*{A. Juridification}

Whether Meta’s normative system can be qualified as law is problematic since it is a private ordering outside the laws of the state world. In Teubner’s framework, social norms coalesce into non-state legal systems when three conditions are met: hierarchization, temporalization, and externalization.\textsuperscript{26} Indeed, non-state normative systems are based on a paradoxical self-validation. They ‘hide’ the paradox, first, by establishing an internal hierarchy of contractual rules. Second, these rules “temporalise the paradox and transform the circularity of contractual self-validation into an iterative process of legal acts, into a sequence of the recursive mutual constitution of legal acts and legal structures.”\textsuperscript{27} Third, contractual rules externalize their self-validation “by referring conditions of validity and future conflicts to external ‘non-contractual’ institutions which are nevertheless

\begin{flushleft}
(describing DNS as a new regulatory space opened up by digital technologies); Christoph B. Graber, \textit{Bottom-Up Constitutionalism: The Case of Net Neutrality}, 7 \textsc{Transnat’l Legal Theory} 524 (2016) (describing how a theory of internet governance—net neutrality—is being promoted and shaped by different social groups on social media platforms); Lee A. Bygrave, \textit{Internet Governance by Contract} 85–103 (2015) (describing how social media platforms such as Facebook are creating new forms of regulatory normativity by seeking to regulate themselves); Christopher Kuner, \textit{Transborder Data Flows and Data Privacy Law} 160 (2013) (explaining the need for a new and more comprehensive framework for regulating transborder data flows); see Vagios Karavas, \textit{Governance of Virtual Worlds and the Quest for a Digital Constitution, in Governance of Digital Game Environments and Cultural Diversity: Transdisciplinary Enquiries} (Christoph B. Graber & Mira Burri-Nenova eds., 2010).


\textsuperscript{26} See Teubner, \textit{Breaking Frames}, supra note 24, at 212. For an empirical account, see Oren Perez, \textit{Transnational Networks and the Construction of Global Law, in The Oxford Handbook of Global Legal Pluralism} 473–489 (Paul Schiff Berman ed. 2020). To be sure, and although the infrastructural, technical and social layers of the Internet remain distinct, similar processes can also be individuated in the normative systems turning around the ICANN or blockchain technologies.

\textsuperscript{27} Teubner, \textit{Breaking Frames}, supra note 24, at 211–12.
\end{flushleft}
'contractual' since they are a sheer internal product of the contract itself.”

How does this suit Meta’s emerging normative system? Meta’s decisions on content moderation are based on a relatively stable hierarchy between primary and secondary norms in the Hartian sense. The Standards of the Community and the Values, the Charter and the Bylaws of the OB provide a quite elaborate, multi-layered set of rules of recognition, adjudication, and change. Such rules govern content moderation decisions in both automated and human review processes, at both substantive and procedural levels. Furthermore, from a temporal point of view, this process is not static. Rather, it is dynamic, as it involves the iterative production of decisions based on existing rules which in turn inform the production of new rules and the evolutive interpretation/implementation of existing ones. The decisions of the OB have precedential value and—especially through the policy advisory statements—contribute to the reflexive adaptation of Meta’s content moderation practices. Finally, such process has been realized through contract rules with which Meta committed itself to comply with the decisions of an “external,” independent body—the OB—and with a wide array of both binding and non-binding sources of human rights law

28. Id. at 212.
33. For a description of the hybrid (AI/human) Facebook internal compliance system, see Perez & Wimer, supra note 8.
34. OVERSIGHT BOARD, supra note 31, at Art. 2.2.
35. Id. at Art. 3.4.
36. “For example, the ‘Dangerous individuals and organizations’ policy was changed following FB-P93JPX02, Case decision 2021-009-FB-UA (Shared Al Jazeera post case).” Contrary to the decisions concerning single pieces of content, policy advisory statements issued by the OB are not binding on Meta and are considered as recommendations. However, Meta has the obligation to publicly disclose the action it takes in response to them. See OVERSIGHT BOARD, supra note 37, at Art. 3.4, Art. 4; OVERSIGHT BOARD, supra note 32, at Art. 2.2. See generally META, Oversight Board Recommendations, (Mar. 17, 2023), https://transparency.fb.com/en-gb/oversight/oversight-board-recommendations/. An overall assessment can be found in Douek, supra note 2.
37. Notably a Trust Agreement and a Limited Liability Agreement, which formally incorporated the OB as an LLC. See generally Klonick, supra note 2, spec. 2467-2469. As such, the OB is a legal entity distinct from Meta, structured as a “non-charitable purpose trust” under Delaware law. An Oversight Board Trust has been in existence since 16 October 2019, but it should not be confused with the OB itself.
(HRL) and codes of conduct. In this sense, the case law of the OB may contribute to the development and conceptualization of an internal distinction within Meta’s normative system, that is, between its “constitutional” and “administrative” rules.

The three dimensions of Meta’s juridification—hierarchization, temporalization, externalization—are particularly apparent in decisions on human rights. The only human right recalled as a parameter of substantive adjudication in the OB’s founding instruments (Charter and Bylaws) was the “fundamental” freedom of expression. The first expansion occurred with the adoption by the OB, in November 2020, of its Rulebook for Case Review and Policy Guidance. This document provided a slight but crucial re-definition of the OB’s own mandate and referenced the Guiding Principles on Business and Human Rights (UNGP) endorsed in 2011 by the UN Human Rights Council. Based on

39. Understood as the set of normative relationships between, on the one hand, OB’s decisions and Meta’s primary norms concerning content moderation and, on the other hand, the higher, overreaching normative layer: e.g., the adequacy of the formulation of a specific community standard with the values or the relevant human right.
40. Understood as the set of normative relationship between, on the one hand, OB’s decisions and Meta’s primary norms concerning content moderation; and, on the other hand, the concrete measures as implemented by automated and human moderators.
41. “Freedom of expression is a fundamental human right . . . The purpose of the board is to protect free expression by making principled, independent decisions about important pieces of content and by issuing policy Advisory Opinions on Facebook’s content policies.” See OVERSIGHT BOARD, supra note 31, at 3. “When reviewing decisions, the board will pay particular attention to the impact of removing content in light of human rights norms protecting free expression.” Id. at Art. 2.2. In the OB Bylaws, human rights are not referred to as a parameter of adjudication, but only in relation to the notification to the participation of the members of the OB in training on Facebook’s policies, values, enforcement practices, and international human rights standards (OVERSIGHT BOARD, supra note 32, at Art. 1§1.4.4), to the expertise of staff members within the OB’s administration (Id. at Art. 1 §2.2), the analysis of how the board’s decisions have considered or tracked the international human rights implicated by a case in the OB’s annual report (Id. at Art. 1 §4.1), and to the principles guiding the notice the posting and referring persons of OB’s decisions (Id. at Art. 2 §2.3.2).
43. See Oversight Board Rulebook for Case Review and Policy Guidance: “The Oversight Board was created to make principled, independent, and binding decisions on what content Facebook and Instagram should allow or remove, based on respect for freedom of expression and human rights [emphasis added].” Note how the language changes compared to the founding instruments. See supra note 42.
44. OVERSIGHT BD., RULEBOOK FOR CASE REVIEW AND POLICY GUIDANCE 3 (2020) (“In line with the UN Guiding Principles on Business and Human Rights, the Rulebook provides greater transparency to users and the general public regarding the function of the Board.”). For a contract-law perspective’s critical take on this development, see Mårten Schultz, Six Problems with Facebook’s Oversight Board. Not Enough Contract
that, the OB, starting from the first decisions published on 28 January 2021, included the UNGP among the “relevant standards” considered for the purposes of its decisions.45 Through the UNGP, the OB considered as “relevant standards” a wide array of sources of HRL mostly directed to states,46 such as the International Covenant on Civil and Political Rights (ICCPR), the International Covenant on Economic, Social and Cultural Rights (ICESCR), and the Convention on the Rights of the Child (CRC), but also texts such as General Comments of Human Rights Expert Bodies and Reports of UN Special Rapporteurs. A few weeks later, in March 2021, Meta announced its new Corporate Human Rights Policy, where the company officially recommitted47 to respecting human rights as set out in the UNGP and in several international instruments.48 This move has only increased the recourse to HRL by the OB, which now uses Meta’s Corporate Human Rights Policy as a “legal” basis to apply human rights standards.

This outline showed how societal constitutionalism interprets Meta’s juridification and the role played by the OB within it. At the same time, it showed the limits of such a process. The latter only covers Meta’s decisions related to content moderation. Furthermore, some significant instruments governing the decision-making process—such as the Internal Implementation Standards49 and the code informing AI-driven content skimming—are not publicly available and cannot orient the conduct of the users as their final addressees. These insights offer a bridge to the next section, which looks at the constitutionalization of Meta.


46. For the limits coming from this HRL feature, see Andreas Kulick, Meta’s Oversight Board and Beyond – Corporations as Interpreters and Adjudicators of International Human Rights, 22 L. & Prac. Int’l Cts. & Tribunals 161, 167–70, 179–80, 192–93 (2023) (arguing that corporations are flawed HRL interpreters but that no alternative to such interpretation exists); Evelyn Douek, The Limits of International Law in Content Moderation, 6 U.C. IRVINE J. INT’L, TRANSNAT’L & COMPAR. L. 37, 50–66 (2021) (discussing seven weaknesses as reasons why HRL will not constrain social media platforms’ operations). See generally Yuval Shany, Digital Rights and the Outer Limits of International Human Rights Law, GERMAN L.J. (forthcoming) (analyzing how “three generations” of digital human rights law fit within the already-changing nature of HRL).


49. Instructions created to enable high-speed micro-deliberation by human moderators.
B. Constitutionalization

Societal constitutionalism distinguishes between juridification and constitutionalization. For the purposes of constitutionalization, a stable distinction between primary and secondary norms is not sufficient. The fact that legal norms govern the recognition, change, and adjudication of other legal norms within a system—put differently, the emergence of legal reflexivity—is only part of the story. Equally necessary is that such legal reflexivity is coupled to the social reflexivity of the involved organization or system, that is, the fact that the organization or system is subject to the operations it produces. Constitutions only arise when phenomena of double reflexivity emerge—reflexivity of the social system that constitutes itself and reflexivity of the law that supports this self-constitutionalization. Constitutions are sociolegal phenomena whereby legal reflexivity is structurally coupled to the social reflexivity of a given system/organization—be it a state, a corporation, or other. In this context, societal constitutionalism aims to capture in legal-theoretical terms the processes that take place in a system to verify whether and to what extent they are constituted and limited by legal norms.

Where does Meta stand in this picture? To answer this question, one must look at the core constitutional features individuated by societal constitutionalism: processes, arenas, and functions.

Meta’s system presents constitutional processes, as it is reflexively subject to producing its own operations/decisions, which are increasingly concerned with the discourses on general constitutional principles that should apply in the digital space, especially social networks, and the

---


51. Introduced in systems theory to explain intersystem relationships, the concept of “structural coupling” indicates a situation in which “a system presupposes certain features of its environment on an ongoing basis and relies on them structurally [. and] . . . the forms of a structural coupling reduce and so facilitate influences of the environment on the system.” NIKLAS LUHMANN, LAW AS A SOCIAL SYSTEM 382 (Fatima Kastner et al., Klaus A. Ziegert trans., Oxford University Press 2004) (emphasis omitted). It emphasizes the constant possibility that systems must link their respective structures in certain situations without losing their identity. Structurally coupled systems thus share some structural elements or some bases of meaning from which, however, they derive different and independent information [through/, which then enters] their respectively different communicative processes.

52. See generally TEUBNER, supra note 12, at 73–110 (describing the roles of these core constitutional features as requirements of a material constitution).
question of constitutional controls. By now, such operations/decisions—at least in the limited content-moderation field—are coupled with the legal reflexivity emerging from the Standards of the Community and Values complex, as the OB interpreted and developed it. In other words, Meta’s normative system presents some limited but stable structural coupling between its social and legal reflexivity. In this sense, the OB is contributing to Meta’s constitutionalization in at least two ways: Firstly, the OB is developing and catalyzing new—substantive and procedural—rules concerning content moderation. Secondly, the OB is using its leverage to ask questions and obtain information about the most untransparent parts of Meta’s system, giving users, regulators, and affected outsiders the opportunity to complain and pursue both contestatory and regulatory strategies.

This last point relates to the emergence of constitutional arenas. Undoubtedly, a clear divide has emerged between organized/decisional arenas and spontaneous arenas, with its distinctive power struggles and contestation practices (i.e., Meta’s corporate governance and moderators versus users and affected outsiders). In this context, societal constitutionalism conceptualizes democratization as the institutionalization of a system’s self-contestation. The never-ending series of scandals and abuses that emerged in recent years—the Cambridge Analytica and the Myanmar scandals in 2018 and the “Facebook files” leak in 2021 are only the tip of the iceberg—sparked global backlash and contestation, boycott campaigns, substantial value

53. More precisely, as far as social reflexivity is concerned, one could distinguish between (1) algorithmic reflexivity—that is, the hierarchy of the bit-by-bit constitution’s different digital levels, in which the highest level is the digital code’s constitutional principles formulated in the program language; and (2) the “political” discourse on general constitutional principles that should apply in the digital space, especially social networks, and the constitutional-controls question. One could also think of algorithmic reflexivity as a quasi-legal reflexivity in which legal constitutional norms are translated into digital operations and programs. I would like to acknowledge that this observation comes from interactions with Nofar Sheffi as well as engagement with her (forthcoming) work.

54. But for important reflections on how Big Tech’s current profitability crisis affects their openness to social, non-profit-driven pressures and their overall accountability, see Kate Klonick, The End of the Golden Age of Tech Accountability, THE KTHONICKLES (Mar. 3, 2023), https://klonick.substack.com/p/the-end-of-the-golden-age-of-tech; see also infra Section 4 (suggesting that social programs do not align with profitability, so the two may be positively correlated such that a decrease in profits leads to a decrease in programs).

55. Cf. sources cited supra note 36 (discussing Oversight Board’s ability to publicly recommend changes to Meta’s policies). Here, I understand “contestatory” in the sense of Philip Pettit, Democracy, Electoral and Contestatory, 42 NOMOS 105 (2000).

loss on the equity market, legal sanctions, as well as (threats of) public regulation, coming from a multitude of individuals, movements, as well as economic and political actors. Meta’s constitutionalization has so far mostly taken the form of a reflexive response to massive, sustained, and ever-growing pressures from its social environment rather than from actual, hard-law obligations or often-timid sanctions by national bodies. By voluntarily—but not spontaneously—binding itself to the decisions of an independent adjudicator, which tends to be sensitive to external social impulses, Meta channels social energies and internally stabilizes the possibility of contestation. By these means, Meta aims to legitimize itself while preserving its operational (and business) autonomy. Here again, societal constitutionalism has explored the constitutionalizing potential of the interaction between “public” and “private” codes of conduct—emerging in the UNGP’s de facto incorporation into Meta’s normative system. All this notwithstanding, and despite these emerging constitutional elements, Meta’s constitutionalization is still far from effective, even from the societal-constitutionalism perspective. Indeed, the key elements are the constitutional functions. According to societal constitutionalism, at the functional level, a constitution emerges if and insofar as the system’s legal norms perform both constitutive/foundational functions and limiting functions toward the dynamics generated by the system’s operations or, more generally, the expansive dynamics that might endanger a system’s social environment and, ultimately, the system itself. The emergence of a normative hierarchy coupled with Meta’s operations, a (quasi-)judicial review, and constitutional arenas points to a constitutive/foundational dimension. Such developments certainly contribute to formalizing and autonomizing Meta’s system by legal means as well as building its symbolic narrative, enshrined in its (quite bombastic) mission statement.

The limiting function concerns limiting Meta’s colonizing tendencies. Here, it is not enough to focus on the crucial but still relatively narrow content-moderation issue. Once again, the scope of the


58. See generally Teubner, Self-Constitutionalizing TNCs? On the Linkage of “Private” and “Public” Corporate Codes of Conduct, supra note 25 (arguing that the interaction between public and private codes of conduct marks the beginning of a constitution and that this process is centered in society).

OB’s “jurisdiction” is limited to the disputed takedowns of single posts or comments. To be sure, the relevance of filtering fake news, hate speech, cyberbullying, and cyberstalking, and thus deciding the limits of free speech of no less than 2.5 billion monthly active users worldwide—often in cases related to extremely sensitive and consequential issues in national and global politics—cannot be underestimated. Still, this is only the most visible side of a much subtler, broader, and to some extent de-personified societal process—in societal constitutionalism’s terms, an “anonymous matrix”.60 Indeed, what comes to the surface and reaches the general public are mostly decisions on content—that is, the what of communication. However, even more problematic, and ostensibly outside the OB’s radar, are the social consequences and negative externalities of social-media-based communication—that is, the how of communication. Social media companies like Meta have a business model centered around monetizing and maximizing attention, as their profits mostly come from advertisers resorting in different ways to online platforms to increase their advertisement success rate. To maximize profits, such companies collect and process as many (meta)data as possible, with the result that a veritable ubiquitous commercial surveillance has emerged. This model is intrinsically oriented toward the maximization of users’ engagement and, to that purpose, resorts to habit-forming digital products, algorithms, and persuasive technologies.61 The engagement growth compulsion pushes single individuals, collective actors (political movements; business entities; media, education, research institutions, etc.), and ultimately

60. See Gunther Teubner, The Anonymous Matrix: Human Rights Violations by ‘Private’ Transnational Actors, 69 MOD. L. REV. 327, 340–42 (2006). Simon Caney has expressed this concept with the term “global injustice.” See Simon Caney, Responding to Global Injustice: On the Right of Resistance, 32 SOC. PHIL. & POL’Y 51, 59 (2015) (“[G]lobal injustice . . . can occur even where there is not necessarily a single clearly defined unified agent behind the injustice. It can occur in cases characterized by three features: (a) a person lacks his or her entitlement (b) this is brought about by a collective causal process; that is, the lack of entitlement is the result of acts by very many (possibly billions) of people. (c) the lack of entitlement is overdetermined; that is, it will occur whenever a certain number of acts are performed. The acts of each are, therefore, individually unnecessary to cause the lack of entitlement.”).

61. See Nir Eyal with Ryan Hoover, Hooked: How to Build Habit-Forming Products 2, 6–13 (Penguin Books 2014); see also B.J. Fogg, Persuasive Technology: Using Computers to Change What We Think and Do 1–2, 6, 8 (2003) (discussing companies’ use of technology that is interactive and stores/manipulates data, both of which increase the technology’s persuasiveness). See generally Nick Lawrence, UI/UX Design: Creating Addictive Products, MEDIUM: UX PLANET (July 2, 2020), https://uxplanet.org/ui-ux-creating-addictive-products-6bfa571c916a (instructing the reader on four-step process to create addictive/habit-forming products).
Meta itself to develop forms of addiction. This economic model, based on the extraction of value from human attention, has probably been best described by Zuboff as “surveillance capitalism.”

Entities like Meta—and, more generally, the “datafication” of governance—do not just shape individual behavior but also drain other social sectors of their functional autonomy. They risk becoming societal black holes as they slowly but inexorably capture other systems in their ever-growing accretion disc. Because of scale and pervasiveness—in terms of number of users, data collection capacity, and manipulation capability—they trigger vicious circles with potentially highly negative impacts on individuals and societal processes alike. The potentially negative effects that social networks can have on the well-being of children, teenagers, and adults, also depends on their social status, wealth, political engagement, and social fragmentation are by now documented and have reached the broader public.

To be sure, social networks should not be demonized and have many positive aspects: increased possibilities of connection, opening of new venues of participation, innovative modalities of social and economic interaction, circulation of information and knowledge, and inclusion of marginalized groups and peripheral communities. Further, several


63. See generally Fleur Johns, Governance by Data, 17 ANN. REV. L. & SOC. SCI. 53 (2021) (surveying the effects of governance by data on states, subjects, and populations).

64. See THE PHILOSOPHY OF ONLINE MANIPULATION (Fleur Jongepier & Michael Klenk, eds., 2022).


67. Digital activist and ethical designer Tristan Harris has coined the phrase “human downgrading” to describe a system of mutually reinforcing harms: addiction, distraction, isolation, polarization, fake news: see Center for Humane Technology, How Technology is “Downgrading Humans” (Tristan Harris X Capgemini), YOUTUBE (Feb. 3, 2022) https://www.youtube.com/watch?v=LZ0PnUzRh8U.

social media companies are not based on that business model. The dangers posed by entities like Meta—the social dynamics to be limited by their “constitution,” if any—do not lie in the ever-increasing interconnection as such. Rather, it is the manipulation of individuals and the related social effects deriving from a business model based on the compulsion towards the maximization of engagement.  

Back to Meta’s normative system, the social dynamics triggered by the uncontrolled maximization of engagement are not addressed, let alone limited by them. Put differently, even within societal constitutionalism’s theoretical framework, Meta’s normative system is well behind in any process of constitutionalization because it does not perform any significant limiting function towards its most disruptive social dynamics. To make a comparison with state-centered constitutionalism, it is like a set of “constitutional” texts addressing the issue of abuses against slaves instead of slavery itself.

Any attempt to establish an effective constitutionalization in the digital sphere must target the vicious circle of monetization of engagement and compulsion to social manipulation. A strong intervention in the form of regulation by states and political actors is necessary. Indeed, even in societal constitutionalism’s legal pluralist framework, state law and constitutions remain central. But targeting the micro-social effects of social media is difficult even for well-motivated and informed public regulators. The problems raised by informational capitalism and so-called human downgrading will be at the forefront for years and call for institutional and regulatory imagination. If the goal is to regulate and constitutionalize social media networks while still retaining their advantages, constitutionalizing impulses need to be effectively internalized into the decision-making premises of the actors involved. However, under current conditions, it seems unlikely that limits to the monetization of attention will be established by Meta’s internalization of external social pressures. Given the limited material scope of its jurisdiction, the OB seems largely powerless. Or maybe not completely.

III. SURVEILLANCE CAPITALISM ON TRIAL?

This section aims to provide a sort of script to build an OB case on the mental health effects of social media, more specifically, on the harmful effects deriving from body comparison and appearance-related

---

69. See THE PHIL OF ONLINE MANIPULATION, supra note 64.
70. See Center for Humane Technology, supra note 67.
The Transformative Potential of Meta's Oversight

Admittedly, it is a hard test for the constitutionalizing potential of the OB, because these effects are not necessarily related to the single pieces of content as such but to the continuous, iterative publication, sharing, and exposure to idealized body images and, more generally, objectifying media, especially for young women. To be sure, such outline cannot provide “the” strategy. Rather, it aims to start a conversation on how to strategically thematize de-personalized, capillary social phenomena in constitutional terms with some of the instruments available. This strategy lies at the core of the practice of public interest litigation, that is, selecting and bringing a case before adjudication bodies with the aim to create broader societal changes. As is known, such practice has historically been one of the most relevant instruments in twentieth-century constitutionalism, especially in the systems of common law tradition.

Outlining such a strategy requires thinking about the preparation and reporting (3.1), the authority of the OB and the scope of its decision (3.2), the relevant standards (3.3), the content of the decision, policy advisory statements, and implementation (3.4).

A. Preparing the Case

On the issue of young users’ mental health, it seems unlikely that Meta will submit to the OB a case for direct review or a request for policy guidance. Therefore, the case should start as a user-generated appeal against a decision made by Meta’s moderators over a single piece of content.

71. This effort does not necessarily coincide with the no less important but different issue of targeted advertising on social networks, which is increasingly entering the radar of regulators. 2022 O.J. (L 277) 27 (prohibiting ad profiling of minors in Art. 28, ¶ 2 and discussing the use sensitive personal data (such as racial or ethnic origin, political or religious affiliation, sexuality or health data) for behavioral targeting in Art. 26 ¶ 3).


75. See Teubner, supra note 60.

76. OVERSIGHT BOARD, OVERSIGHT BOARD BYLAWS, Art. 2 §§ 1, 2.1, 2.1.2 (2022).

77. Id. at Art. 2 § 1; Id. at Art. 3 § 1.2; Id. at Art. 2 § 2.1.3.
Given that the OB has the discretion to choose which requests it will review and decide upon, thinking strategically about such a case requires the involvement of public opinion and institutional actors. The latter should be involved in a coordinated effort to create pressure, expectations, and social oversight on the OB. In other words, the initiative should set the conditions for the case to be selected by the OB under the Charter provisions and its own “Overarching Criteria for Case Selection.” According to the latter, the OB Case Selection Committee—which decides by a majority vote—selects cases for review that raise “important issues pertaining to respect for freedom of expression and other human rights and/or the implementation of Facebook’s Community Standards and Values” and are “of critical importance to public discourse, directly or indirectly affect a substantial number of individuals, and/or raise questions about Facebook’s policies.”

Against this background, the case should be preceded by a broad campaign, receiving as much public coverage as possible, by both institutional and civil society actors, and targeting specifically the OB. Ideally, the appeal should be endorsed and covered not only by digital activists and organizations institutionally involved on such issues, but also by policy-makers across the political spectrum not necessarily invested in new technology issues, (human rights) NGOs, press, high-profile officials, national and international institutional bodies such as the UN Special Rapporteur on the Right to Health and the World Medical Association.

To increase the chances for the case to be selected, the initial report

---

78. Id. at Art. 2 § 1.
79. Id. at Art. 2, § 1; “In its selection, the board will seek to consider cases that have the greatest potential to guide future decisions and policies” (emphasis added).
80. Id. at Art. 1 § 3.1.2.
81. The updated criteria are available at: Oversight Board, Overarching Criteria for Case Selection, https://oversightboard.com/sr/overarching-criteria-for-case-selection. The Rulebook for Case Review and Policy Guidance describes this procedure as follows: “Step 1: Selection of Cases for Manual Review. 1) Based on the overarching selection criteria established by the Oversight Board and specific selection criteria established by the Case Selection Committee, the Case Selection Team will filter user-generated appeals and Facebook-referred cases in the CMT. 2) The Case Selection Team will manually review the filtered cases to present a reviewable longlist to the Case Selection Committee, alongside statistics about both the longlist and non-selected cases. 3) The Case Selection Committee will identify a shortlist of cases from the longlist that it will consider for selection. The Case Selection Committee may review and adjust the specific criteria if they wish to include more cases.”
should ideally be filed by multiple young users with active Facebook or Instagram accounts who have experienced or are experiencing mental health distress in relation to self-esteem, body dissatisfaction, and appearance-related social media consciousness. Even without leading to severe forms of depression or self-harm, such distress should have made them stop or reduce their use of social media, notably Instagram. Ideally, such reduction of the use should follow recommendations provided by medical professionals, in turn grounded on the growing evidence on the correlation between the emergence of mental health distress in young people and the use of social media. At the same time, social media should play a particularly relevant role in the social interactions of the reporting users, for example, because, due to geographical, economic, or social reasons, they cannot easily access alternative sources of social interaction, information, or inclusion.

The initial report should target any piece of content—ideally a body picture on Instagram—which, while not necessarily promoting a specific standard of unattainable beauty, may however reasonably induce feelings or thoughts of body inadequacy/dissatisfaction, especially among teenage girls. The piece should also be a particularly “successful” one, possibly created or shared by social media influencers, highly followed accounts, or other ‘public’ figures, reaching high virality, for example in the form of a “meme.” In other words, the reporting strategy should focus on pieces of content that, while in the abstract “harmless” per se, might nonetheless, by their spread and iterative presence on a social media platform, reasonably generate or reinforce broader harmful mental health patterns among most vulnerable users. In other words, the report (and the subsequent appeals to Meta’s reviewers and then to the OB) should link as far as possible the actual—possibly non-harmful—piece of content to the harm caused by the cumulative effects of its widespread diffusion.

Leaving aside for now the content policy upon which the report


84. See also supra note 9; Haidt, supra note 10.

85. OB Bylaws, Art. 2, Sec. 1.2.1 (demonstrating that the report should not target the following types of content, explicitly excluded from review which includes content posted through marketplace, fundraisers, Facebook dating, messages, and spam).

86. Paul Gil, What is a Meme?, LIFEWIRE, Sept. 16, 2022, https://www.lifewire.com/what-is-a-meme-2483702 (defining a meme as ‘a virally transmitted image embellished with text, usually sharing pointed commentary on cultural symbols, social ideas, or current events. A meme is typically a photo or video, although sometimes it can be a block of text’).
could be based,\textsuperscript{87} it is improbable that automated or human reviewers will take down the reported piece of content. First, Meta’s content policies and the OB’s rules target individual pieces of content\textsuperscript{88} that (may) \textit{directly} cause harm. Second, even if a clear “legal” basis or cause of action was available, the nature of the harm—primarily deriving from the \textit{cumulative} effects of individual pieces of content—makes it extremely difficult for moderators to single out pieces to take down. Mutatis mutandis, this kind of litigation faces some of the difficulties of climate corporate litigation when it targets emissions of private actors whose actual contribution to climate change is difficult if not impossible to individuate.\textsuperscript{89}

Therefore, the reporting user(s) would see their initial request and subsequent “administrative” appeals\textsuperscript{90} rejected by Meta’s moderators. At that point, once the “administrative” procedure is exhausted, they may finally submit a request for review to the OB, starting the “judicial” procedure.\textsuperscript{91} Importantly, when submitting the case to the OB, the users should fully exploit, ideally with the assistance of digital activists, digital design ethicists, and psycho-medical professionals, the opportunity provided by Article 3, Section 1.2.1 of the Bylaws, whereby reporting users may “explain why they believe [Meta] may have made an incorrect decision on the content in question; why they believe the board should hear their case; why they originally posted the content; and how [Meta]’s decision could impact others.” The submission, in other words, should be carefully crafted to, first, request a content moderation decision; and second, fully reflect the public interest significance of the case,\textsuperscript{92} possibly incorporating the best scientific knowledge available.

\begin{itemize}
\item \textsuperscript{87}OB Bylaws, Art. 3, Sec. 1.3, OVERSIGHT BOARD, 29 (Feb. 2023), \url{https://www.oversightboard.com/sr/governance/bylaws} (this issue is addressed also in Sec. 3.3).
\item \textsuperscript{88}OB Bylaws, Art. 3, Sec. 1.1.1, OVERSIGHT BOARD, 27 (Feb. 2023) \url{https://www.oversightboard.com/sr/governance/bylaws}.
\item \textsuperscript{89}See also Caney, supra note 60.
\item \textsuperscript{90}See Klonick, supra note 37 at 1647–48;Douek, supra note 20 at 563–64.
\item \textsuperscript{91}OB Bylaws, Art. 3, Sec. 1, Sec. 1.1, OVERSIGHT BOARD, 27 (Feb. 2023) \url{https://www.oversightboard.com/sr/governance/bylaws} (explaining these requests must be made within fifteen days of the final decision).
\item \textsuperscript{92}OB Bylaws, Art. 2, Sec. 2.1.1, OVERSIGHT BOARD, 22 (Feb. 2023) \url{https://www.oversightboard.com/sr/governance/bylaws} (OB Bylaws concerning the submission by Meta of ‘significant and difficult’ cases: ‘Significant means that the content in question involves real-world impact and issues that are severe, large-scale, and/or important for public discourse. Difficult means the content raises questions about current policies or their enforcement, with strong arguments on both sides for either removing or leaving up the content under review’).
\end{itemize}
B. Case Selection: Authority and Scope

When it comes to the authority and scope in such a case, the OB has a clear power to review Meta’s decision following an appeal from the person who previously submitted the content for review. At the case selection stage, however, Meta still has the power to review the shortlist of cases identified by the OB Case Selection Committee before a final decision over the selection is made. Indeed, according to the Bylaws, “[n]othing in these bylaws or other governing documents will be interpreted in a manner that would result in a violation of law by [Meta]” and, for that purpose, cases that meet specific criteria are not eligible for the OB to review.

However, Meta would have no ground to find the case we are outlining here as non-eligible. Indeed, as the appeal aims to take down content upheld by Meta’s moderators after the initial report, most ineligibility criteria listed in the Bylaws would not be met. First, the decision of the OB would not concern a piece of content that has already been blocked, following the receipt of a valid report of illegality, and not removed for a Community Standards violation. Second, by design, the case does not concern content that is “criminally unlawful in a jurisdiction with a connection to the content . . . and where a OB decision to allow the content on the platform could lead to criminal liability for [Meta], [Meta] employees, the administration, or the [OB]’s members.” Third, it is doubtful that any OB decision to take down the content of the sort discussed here could lead to “adverse governmental

---


94. See Oversight Board Criteria for Case Selection, supra note 81 (the shortlist will be sent to Facebook for legal review. Facebook may exclude from the shortlist cases that are not eligible for review by the Oversight Board, in accordance with the Bylaws, such as cases that could result in criminal liability or adverse government action if reviewed. The Case Selection Team will keep statistics on Facebook exclusions).

95. OVERSIGHT BOARD, The Rulebook for Case Review and Policy Guidance (Nov. 2020), https://oversightboard.com/sr/rulebook-for-case-review-and-policy-guidance (describing this step: ‘The shortlist will be sent to Facebook for legal review. Facebook may exclude from the shortlist cases that are not eligible for review by the Oversight Board, in accordance with the Bylaws, such as cases that could result in criminal liability or adverse government action if reviewed. The Case Selection Team will keep statistics on Facebook exclusions.’)


97. Id.

98. Id.
action” against Meta, Meta’s employees, the administration, or the OB’s members.99

C. Standards of Review

Since its first decisions, the OB has based its review on three distinct sets of standards:100 Values, Standards of the Community,101 and HRL. So far, the OB usually has interpreted these three sets as coherently as possible. However, they are treated as distinct benchmarks in dynamic interaction with each other.

At first sight, the Standards of the Community are designed as a specification/operationalization of the Values. Therefore, the latter appear to inform the application and interpretation of the former.102 This means, for example, that a decision by Meta’s moderators to remove a piece of content may be consistent with the Standards of the Community but still in breach of its Values, thus leading to the restoration of the content by the OB.103

In contrast, Meta’s restoration of content initially removed for a breach of the Standards of the Community may be upheld because the restoration complies with the Values.104 However, looking at the OB’s case law, it is often the relevant HRL that tips the balance in favor of removal or restoration. In this sense, after more than two years of case law, HRL has quietly emerged as the center of gravity in the OB’s review process and the deal-breaker in the most complex cases.105

99. Id.
100. See Meier & Golia, supra note 2 (for an attempt at systemization).
101. See Meta’s Transparency Center, META, https://transparency.fb.com/. See also Case Decision 2020-004-IG-UA (Breast cancer symptoms and nudity case); Case Decision 2021-006-IG-UA (Ocalan’s isolation case); Case Decision 2021-013-IG-UA (ayahuasca case) (including both Facebook Community Standards and Instagram Community Guidelines. Importantly, if content is considered violating on Facebook, it is also considered violating on Instagram).
1. Values

Regarding the Values, a preliminary point concerns the internal relationship among them. “Voice”—the lyrical proxy used to indicate freedom of expression—is described as “paramount.” This formulation suggests a hierarchical superiority or at least a rule-exception relationship with the other Values (authenticity, safety, privacy, and dignity), presented as limitations to an otherwise unrestricted value. If this were the case, the justification for removing content or limiting “voice” would involve an explicitly higher argumentative or evidentiary threshold upon the appellants and, ultimately, the OB. Moreover, in its case law, the OB pays at least some lip service to this apparent rule/exception relationship. However, at a closer look, the actual prevalence between “Voice” and the other (counter-)values in each case is decided less by hierarchy or derogation than by balancing techniques.

This point is crucial for our purposes. However, the question remains on which specific Value may be invoked. Since our case is about removing the targeted content, the four counter-values as potential causes of action come to mind: authenticity, safety, privacy, and dignity. However, their wording is extremely broad, and their normative content is elusive. Further—probably not by chance—they are formulated in such a way to exclude from their scope precisely the type of harm that this litigation seeks to redress: the harm caused by the cumulative effects of a massive number of individual, otherwise non-necessarily harmful pieces of content. In particular, they seem to target the effects of interpersonal communication among users rather than the effects of social media communication as such.

For example, when one thinks that some of the harmful effects of body comparison are due to image filters, the Value of authenticity comes to mind. However, the formulation of the latter (and of the Standards of the Community implementing it) suggests that it only concerns cases of fake or misrepresented identity. When one thinks of some of the most severe forms of mental distress and self-harm

106. See OB Charter, Introduction; see also the Facebook Community Standards (‘Our commitment to expression is paramount, but we recognize that the Internet creates new and increased opportunities for abuse. For these reasons, when we limit expression, we do it in service of one or more of the following values […]’).
108. ‘[W]e don’t want people using Facebook to misrepresent who they are or what they’re doing.’
correlated to the use of social media such as Instagram, the Value of safety comes to mind. But the text clarifies that it only concerns content that may cause \textit{physical} harm or may “intimidate, exclude or silence others.”\(^\text{109}\) Even “dignity” only obliges users not to “harass or degrade others.”\(^\text{110}\)

Therefore, to use the Values strategically, one needs to approach them somehow counterintuitively, that is, focusing on the paramount Value of voice. And indeed, the kind of harm/distress caused to the users leads them or is supposed to lead them to reduce their use of social media, often under the recommendation of medical professionals. This could constitute a cause of action to require the removal of a piece of content. After all, “the goal of [the] Community Standards is to create a place for expression and give people a voice” and, if one takes Meta’s statements seriously, its mission is “to give people the power to build community and bring the world closer together.”\(^\text{111}\) But what voice is left to users who must reduce their use of Meta’s platforms? Were they not pushed farther away from the world—under Meta’s vision at least?

To be sure, this would be a rather formalistic argument. However, it has the advantage of turning Meta’s most powerful rhetoric against it. Further, pushing Meta to somehow reduce the presence of certain types of content and body pictures—especially when linked to habit-forming digital products—might decrease users’ engagement, leading to a loss of profitability and, ultimately, of the market value of Meta itself.\(^\text{112}\) This, however, would be a business-related issue that could not be invoked as a counterargument before the OB. The latter is an independent body that applies the rules of Meta’s \textit{normative} system.

\section*{2. Standards of the Community}

Turning to the community standards, the same problems encountered when dealing with Values come back here. First, their formulation makes finding a solid cause of action challenging. Second, the targeted harm is generally conceived as directly attributable to a single piece of content posted by another user or account. The “danger,” for example, is always perceived as coming from single individuals or

\begin{itemize}
  \item \textit{109.} The full text reads: ‘[W]e’re committed to making Facebook a safe place. We remove content that could contribute to a risk of harm to the physical security of persons. Content that threatens people has the potential to intimidate, exclude or silence others and isn’t allowed on Facebook.’
  \item \textit{110.} The full text reads: ‘We believe that all people are equal in dignity and rights. We expect that people will respect the dignity of others and not harass or degrade others.’
  \item \textit{111.} \textit{See} Meta Investor Relations, \textit{supra} note 59.
  \item \textit{112.} \textit{See infra, Section 4.}
\end{itemize}
organizations, never from the features of the communication as such. In other words, and once again, it is extremely difficult to target harms caused by the cumulative effects of a massive number of otherwise non-harmful pieces of content. This is reflected in the “immediate harm” standard applied by the OB in several cases concerning hate speech, incitement to violence, dangerous individuals and organizations, and regulated goods.

However, it still seems possible to individuate some “hooks” to make the systemic effects enter the OB’s review process and cross the thin line between non-harmful pieces of content and their broader societal impact.

Firstly, both the Charter and the Bylaws refer to the possibility of taking action against “identical content with parallel context” of a piece of content that has been removed following a binding decision of the OB. This means that even Meta recognizes at some level that the impact of the OB making decisions extends beyond the content in the single case and must be considered by Meta itself. Secondly, in stressing that “context matters,” the OB has repeatedly specified that it matters both online and offline. This means that broader, real-life,

113. Variously indicated also as ‘imminent harm’ or ‘immediate risk of harm’.
117. See Ayahuasca Brew Case Decision; see also Asking for Adderall® Case Decision 2021-015-FB-UA Facebook Oversight Bd., https://www.oversightboard.com/decision/FB-Q72FD6YL.
118. OB CHARTER Art. 4.
119. OB BYLAWS, Art. 2, Sec. 2.3.1.
120. See Breast Cancer Symptoms and Nudity Case Decision; see also Myanmar Bot Case decision 2021-007-FB-UA Facebook Oversight Bd., https://oversightboard.com /decision/FB-ZWQUPZLZ/.
and—importantly—only potential consequences of a single piece of content play a role in the decision over its removal or restoration. Thirdly and relatedly, some decisions of the OB open a promising window to take into consideration the systemic effects of online communication. In this context, at least two cases stand out.

The first case is the Zwarte Piet case. In this case, the OB upheld Meta’s decision to remove a video showing a young child meeting three adults, two portraying the character of Zwarte Piet, with their faces painted black and wearing Afro wigs under hats and renaissance-style clothes. While it is a cultural tradition shared by many Dutch people without apparent racist intent, it includes the use of blackface, recognized as a harmful racial stereotype prohibited by Meta under its “hate speech” policy. Importantly for our purposes, a majority of the OB found that “it cannot be decisive that the user shared this content without malicious intent or hatred towards Black people.” The majority also “noted that repeated negative stereotypes about an already marginalized minority, including in the form of images shared on social media, have a psychological impact on individuals with societal consequences.” Further:

Allowing the accumulation of such posts on Facebook would create a discriminatory environment for Black people that would be degrading and harassing. At scale, the policy is clear and ensures Black people’s dignity, safety and voice on the platform. Restricting the voice of people who share depictions of blackface in contexts

---

9395596088473-oversight-board-overturns-original-facebook-decision-case-2021-009-fb-ua/; Colombia protests Case decision 2021-010-FB-UA https://www.oversightboard.com/news/223462609822963-oversight-board-overturns-facebook-decision-case-2021-010-fb-ua/; FB-TYE2766G, Case Decision 2021-011-FB-UA (South Africa slurs case); Alleged crimes in Raya Kobo Case Decision 2021-014-FB-UA, https://oversightboard.com/news/927673894608838-oversight-board-upholds-meta-s-original-decision-case-2021-014-fb-ua/ (to be sure, the relevance of offline context is not foreign to Meta’s policy rationale concerning hate speech: “We believe that people use their voice and connect more freely when they don’t feel attacked on the basis of who they are. That is why we don’t allow hate speech on Facebook. It creates an environment of intimidation and exclusion, and in some cases may promote offline violence. […] We also prohibit the use of harmful stereotypes, which we define as dehumanizing comparisons that have historically been used to attack, intimidate, or exclude specific groups, and that are often linked with offline violence) (emphasis added); (see also, https://transparency.fb.com/policies/community-standards/hate-speech/).

123. Id. at § 8.2.
124. Id. at sec. 8.3 (emphasis added).
where it is not condemning racism is acceptable to achieve this objective.\textsuperscript{125}

At the same time, when discussing the necessity and proportionality of the removal under human rights standards, the OB held that:

moderating content to address the cumulative harms of hate speech, even where the expression does not directly incite violence or discrimination, can be consistent with Facebook’s human rights responsibilities in certain circumstances. [...] the accumulation of degrading caricatures of Black people on Facebook creates an environment where acts of violence are more likely to be tolerated and reproduce discrimination in a society.\textsuperscript{126}

In other words, for the purposes of the removal, it did not matter whether the content as such was (intended to be) harmful or not. This principle is crucial because for the first time, it delinks the harm of a single piece of content from the harmful effects of its accumulation and spread on the platform.

The second case is the \textit{Alleged Crimes in Raya Kobo} case.\textsuperscript{127} In a highly delicate case concerning an ongoing ethnic conflict, the OB decided for the removal of a post alleging that the Tigray People’s Liberation Front (TPLF) killed and raped women and children, and looted the properties of civilians in Raya Kobo and other towns in Ethiopia’s Amhara region. The poster also claimed that ethnic Tigrayan civilians assisted the TPLF with these atrocities and the information was received from the residents of Raya Kobo. The post ended with the words ‘we will ensure our freedom through our struggle.’ Meta had initially removed the post under its ‘hate speech’ policy. However, after an appeal was submitted to the OB, Meta identified its original decision to remove the post as incorrect and restored it. However, the OB found that the content violated the ‘violence and incitement’ policy. In its reasoning, the OB recalled the ‘accumulating and spreading’ standard

\textsuperscript{125} Id.

\textsuperscript{126} Id. (OB building on FB-QBJDASCV, Case Decision 2020-003-FB-UA (Armenians in Azerbaijan case), specifically concerning situations of armed conflict: “in situations of armed conflict in particular, the risk of hateful, dehumanizing expressions accumulating and spreading on a platform, leading to offline action impacting the right to security of person and potentially life, is especially pronounced.”).

\textsuperscript{127} See, \textit{Alleged crimes in Raya Kobo Case Decision 2021-014-FB-UA, Facebook Oversight Bd.}, https://www.oversightboard.com/decision/FB-MP4ZC4CC [hereinafter Raya Kobo].
introduced by the *Armenians in Azerbaijan* case for situations of armed conflict.\textsuperscript{128} However, this time, it added a crucial specification:

> cumulative impact can amount to causation through a “gradual build-up of effect,” as happened in Rwanda where calls to genocide were repeated […]. A direct call for violence is absent from the post in this case, although there is a reference to “our struggle.”\textsuperscript{129}

These cases are not the only “hooks” in the OB’s case law. For example, in the Trump case, the OB stressed the importance of the “degree of influence that a user has over other users” and that “context matters when assessing issues of causality and the probability and imminence of harm.”\textsuperscript{130}

To be sure, “cumulative effects” or “gradual build-up of effect” are not equivalent to “systemic” effects. Moreover, one must consider that some of these standards have been elaborated for situations of armed conflict or civil unrest. Especially the *Zwarte Piet* case softens the rigidity of the “immediate harm” standard, thus providing a margin of maneuver. With this approach to causation, one may allege violations of the Standards of the Community deriving from even more granular phenomena, e.g., the continuous, iterative exposure to idealized body images. Expanding on such approach to causation, one may find causes of action even in some narrowly formulated policies, particularly “suicide and self-injury”\textsuperscript{131} and “bullying and harassment.”\textsuperscript{132}

### 3. Human Rights

Turning to HRL, there are several norms of both binding and non-binding nature dealing with the right to (mental) health.

---

\textsuperscript{128} See *Zwarte Piet*, supra note 122.


\textsuperscript{131} ‘We remove any content that encourages suicide or self-injury’ (emphasis added).

\textsuperscript{132} “Do not target private individuals […] with: […] - Comparison to other public, fictional or private individuals on the basis of physical appearance. […] - Neutral or positive physical descriptions.” “The Bullying and Harassment policy also seeks to deter behaviour that can cause significant emotional distress and psychological harm, implicating users’right to health” (Pro-Navalny protests in Russia case), § 8.3 (emphasis added).
First, Art. 12, para. 1 ICESCR provides that “The States Parties [...] recognize the right of everyone to the enjoyment of the highest attainable standard of physical and mental health.”

Secondly, more specifically, Arts. 3 and 19 CRC recognize states’ obligation to protect the health and well-being of children. Art. 17 CRC also recognizes the importance of mass media for the “social, spiritual and moral well-being and physical and mental health” of children (para. 1). It obliges States to “[e]ncourage the development of appropriate guidelines for the protection of the child from information and material injurious to his or her well-being” (para. 2, lett. e). According to the authoritative opinion of the Committee on the Rights of the Child, Art. 17 CRC:

delineates the responsibilities of mass media organizations. In the context of health, these can be further expanded to include promoting health and healthy lifestyles among children; providing free advertising space for health promotion; ensuring the privacy and confidentiality of children and adolescents; promoting access to information; not producing communication programmes and material that are harmful to child and general health; and not perpetuating health-related stigma.\(^{133}\)

In the same direction, in a General Comment specifically dedicated to children’s rights in relation to the digital environment, the Committee stated that:

States parties should regulate against known harms and proactively consider emerging research and evidence in the public health sector, to prevent the spread of misinformation and materials and services that may damage children’s mental or physical health. *Measures may also be needed to prevent unhealthy engagement in digital games or social media, such as regulating against digital design that undermines children’s development*

---

133. UN Committee on the Rights of the Child (CRC), General comment No. 15 (2013) on the right of the child to the enjoyment of the highest attainable standard of health (art. 24), (Apr. 17, 2023), CRC/C/GC/15, ¶ 84.
and rights [emphasis added].\textsuperscript{134}

Besides specific HRL instruments, several other international documents of different nature deal with the (mental) health of children in the context of digital technologies and social media.\textsuperscript{135}

To be sure, one of the most challenging issues would be striking the right balance with other rights recognized in multiple instruments of both binding and non-binding HRL, such as free speech, artistic expression, and economic freedom (of business companies and/or individuals such as social media influencers). In its balancing/proportionality exercise however, such rights—however relevant—would recede against the apparent violation of children’s right to (mental) health and, more generally, the protection of the “best interest” of the child.\textsuperscript{136}

Besides the special status of childrens’ rights in international law, a growing consensus is emerging on the necessity to limit and/or prohibit targeted advertising and social profiling of minors. Although such practices do not necessarily cause mental health distress per se, the use of addictive digital technologies on social media platforms is closely linked to the maximization/monetization of attention and the need to collect as much data as possible, as seen above. There is a red thread linking surveillance capitalism as the main business model of social media companies, addictive digital technologies, and mental health distress. This is another reason targeted advertising is increasingly entering the radar of regulators, at least in Europe. For example, the

\textsuperscript{134} UN Committee on the Rights of the Child (CRC), General comment No. 25 (2021) on children’s rights in relation to the digital environment, (Mar. 2, 2021), CRC/C/GC/25, ¶ 96.


\textsuperscript{136} Roberta Ruggero, Article 3: The Best Interest of the Child, in Monitoring State Compliance with the UN Convention on the Rights of the Child 21-29 (Ziba Vaghri, et al. eds., 2022); Jason M. Pobjoy, The Best Interest of the Child Principle as an Independent Source of International Protection, 64 International & Comparative Law Quarterly 327 (2015); UN Committee on the Rights of the Child (CRC), General comment No. 1 14 (2013) on the right of the child to have his or her best interests taken as a primary consideration (art. 3, para. 1), CRC/C/GC/14.
recently adopted Digital Services Act of the European Union\textsuperscript{137} prohibits ad profiling of minors (Art. 28, para. 2)\textsuperscript{138} and the use of sensitive personal data (such as racial or ethnic origin, political or religious affiliation, sexuality, or health data) for behavioral targeting of any user (Art. 26 para. 3). In the same vein, European data protection authorities have recently started fining Meta for the illegal processing of European users’ data—regardless of their age—clarifying that, under Arts. 6 and 21 of the GDPR, “contractual necessity” and “legitimate interest” do not provide sufficient legal bases for processing user data for personalized ads.\textsuperscript{139} As the EU legal system provides legal protections to all the rights involved—economic freedoms, privacy, health, children rights—this development is significant also as a model for potential lines of reasoning of a decision by the OB, and especially for how to conduct the proportionality test amongst conflicting rights.


\textsuperscript{138}. See also recitals 71, 81, 83, and especially 89: ‘Providers . . . should take into account the best interests of minors in taking measures such as adapting the design of their service and their online interface, especially when their services are aimed at minors or predominantly used by them. . . . They should also take measures to protect minors from content that may impair their physical, mental or moral development and provide tools that enable conditional access to such information. In selecting the appropriate mitigation measures, providers can consider, where appropriate, industry best practices, including as established through self-regulatory cooperation, such as codes of conduct, and should take into account the guidelines from the Commission.’ In the same direction, the proposed Artificial Intelligence Act (Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts COM (2021) 206 final, available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206), aims to establish prohibitions in the use of AI covering ‘practices that have a significant potential to manipulate persons through subliminal techniques beyond their consciousness or exploit vulnerabilities of specific vulnerable groups such as children or persons with disabilities in order to materially distort their behaviour in a manner that is likely to cause them or another person psychological or physical harm (emphasis added)’ (5.2.2).

D. Decision

The OB should use this opportunity to at least spotlight the internal functioning and effects of Meta’s content moderation on the mental health of vulnerable people. By exposing the links between the business model, (the lack of) content moderation, and the mental distress of users, the OB would act as a gateway channeling external pressure for (potential) systemic change. Insofar as it facilitates scrutiny of Meta’s internal operations and their societal effects, this move has value in itself. Indeed, it would further expose the company to the global colère publique contributing to its constitutionalization. In this sense, this decision and the OB Advisory Opinion on the “cross-check” could create mutually reinforcing patterns.140

Although the case we imagine would probably originate from an individual complaint rather than Meta’s request for a policy Advisory Opinion, the OB could still “request additional information that is reasonably required for it to make a decision.”141 The information requested should concern (1) whether and to what extent Meta’s algorithms purposefully exposed the user/applicant to the contested piece of content as a consequence of social profiling; (2) whether and to what extent Meta’s algorithms were designed to increase the engagement of the already-distressed and/or addicted user/applicant. While much of this information has already leaked following the Facebook files scandal recalled above, the information obtained by the OB would further broaden current knowledge on Meta’s (algorithmic) decision-making process, especially when it comes to the contribution to users’ mental health distress and its instrumentalization.

However, taking into consideration how unlikely it is that Meta provides this kind of information,142 it would be crucial to receive insights from as many subject-matter experts as possible143—ranging

140. OB, supra note 8.
141. OB Bylaws, art. 2, sec. 2.2.2; see also OB Charter, Art. 2, Sec. 3.
142. See OB Bylaws, art. 2, sec. 2.2.2 (providing reasoning on the grounds for Meta’s refusal to provide information: ‘Facebook may decline such requests where Facebook determines that the information is not reasonably required for decision-making in accordance with the intent of the charter, is not technically feasible to provide, is covered by attorney/client privilege, and/or cannot or should not be provided because of legal, privacy, safety, or data protection restrictions or concerns.’)
143. OB Bylaws, art. 1, sec. 2.1.3: “At their discretion and prior to deliberation, board panels may request and receive information from a global pool of outside subject-matter experts, including academics, linguists, and researchers on a specific issue (e.g., region, cultural norm, or phrase). This pool of experts will be populated at the discretion and/or direction of the board. Separately, panels may, at their discretion, also request issue briefs from advocacy or public interest organizations that reflect a range of perspectives.”
from medical experts to programmers up to digital product designers. This call for public comments should aim not only at “proving” and understanding the links between mental health distress and social media usage in the specific case but also at collecting and systematizing existing knowledge concerning technical solutions and policy options, notably on ethical digital design.

In this sense, the importance of this pooling/systematization of information goes well beyond the decision’s rationale. Even following the best strategies available, the takedown of the contested piece of content remains unlikely. And even in that case, such victory would have a largely symbolic value, given the granular, iterative features of the social media interactions causing mental health distress. However, just like in “analog” public interest litigation, a victory in “court” is only one of the goals, perhaps not even the primary one.\(^{144}\)

Besides the—admittedly unlikely—scenario where the OB decides in favor of the content’s takedown, more relevant—in terms of impact and jurisgenerative effects—is the possibility to issue policy advisory statements in the form of recommendations to Meta. To be sure, recommendations as such are not binding upon Meta. However, they trigger a powerful “comply-or-explain” mechanism,\(^{145}\) further reinforced by the global scrutiny channeled before and after the OB’s decision—and by the OB itself. In other words, this strategy provides a crucial opportunity to showcase a series of concrete technical guidelines and policy options to reduce the addictive effects of social media use and their mental health consequences upon young, vulnerable users.

For this reason, and following again the example of the Advisory Opinion on the “cross-check” program, the OB should not recommend oversimplified solutions such as the—unrealistic—termination of targeted advertising. Rather, based on the information obtained during the procedure, it should provide concrete, realistic guidelines for actions to be taken.

Among them, and short of radical solutions such as the outright elimination of the “like/reaction” button,\(^{146}\) one could include “time spent” indicators, newsfeed filtering, “why I see what I see” disclaimers, notification grouping/muting, and broader and easier access to “do not

---

144. Cf. Harold Koh, Transnational Public Law Litigation, 100 YALE L.J. 2347, 2348 (1991) (referring to the ‘enunciation of a […] norm that will stimulate “relief” in the form of a negotiated political settlement.’)
145. OB Bylaws, art. 1, sec. 2.3 and 2.3.2.
146. Jan Fox, An unlikeable truth: Social media like buttons are designed to be addictive. They’re impacting our ability to think rationally, 47 INDEX ON CENSORSHIP 11,11 (2018).
disturb” settings. These and other changes could be tailored to the vulnerabilities of users according to a broad set of indicators such as background (e.g., refugee), gender, age, and race, among others. The tailoring should take into consideration the specificities of mental health illnesses as well as the fact that there is a significant variety of forms, drivers, and consequences of social media addiction, depending on the type of device, interface, and platform used. Moreover, one has to keep in mind that social networking and social media use are not the same.149

To be sure, strict prohibition rules should prevent such services from becoming a hidden social profiling, to be further monetized. Content moderation must not become a de facto surveillance in itself.150 Rather, users should have the chance to customize their social media experience according to their specific mental health distress and/or illness. Algorithms should modulate individual feeds accordingly, potentially through AI technology pre-identifying triggering content, especially regarding idealized body images and social comparison.151 Crucially, shaping the feed and interface to actively reduce engagement should not be ruled out. Encryption should also be used to keep sensitive data confidential and unavailable for other purposes, including governmental investigation.

Besides the technological specificities, their ecosystem should support both internal openness and external pressures. Internally, such mechanisms—particularly the prohibition of monetization—should be under the supervision of experts who do not report to public policy teams or those in charge of relationship management with affected users.152 Externally, the technology used should be co-designed with

147. See Mujica, supra note 62, at 18–22; and Attila Somos, 7 Ethical Design Examples To Make Facebook Better For Everyone, UXStudio (Apr. 13, 2020), available at: https://uxstudioteam.com/ux-blog/ethical-design/.
151. A related but different issue is the use of AI and deep learning techniques for suicide and depression detection, an option not directly related to content moderation as such which raises different issues, from technological, ethical, and medical perspectives; See Anshu Malhotra & Rajni Jindal, Deep learning techniques for suicide and depression detection from online social media: A scoping review, 130 Applied Soft Computing 109713 (2022).
152. Cf. Meta’s Cross-Check Program Case Decision ¶ 134; see Klonick, supra note 8 (discussing the relevance of this “church-and-state separation.”)
internationally recognized medical associations and public interest groups, representative of affected people.\textsuperscript{153} At the same time, even after their release, such technologies should remain public, transparent, contestable, and subject to constant review\textsuperscript{154} through bottom-up initiatives. In this sense, this kind of policy recommendation would contribute to the effective constitutionalization of content moderation by exploiting the democratizing potential of algorithms and digital co-design.\textsuperscript{155}

These are some suggestions concerning the content of the OB decision, which are neither exhaustive nor “right” in themselves. As said before, this article intends to offer a sort of script; however, it might indicate a potential path to address one of the most sensitive and problematic issues deriving from the widespread use of social media networks within the straitjacket of the rules governing its jurisdiction.\textsuperscript{156}

\section*{IX. Conclusion}

This article imagined a litigation strategy using the instruments available within the normative system of a major social media company to contribute to the effective constitutionalization of digital spaces. The

\begin{itemize}
\item \textsuperscript{154} In this direction, \textit{mutatis mutandis}, see the relevant provisions of the DSA, supra note 142, requiring “very large online platforms” to periodically conduct and publish assessments concerning systemic risks, particularly before launching new services (Art. 34), with mitigation obligations (Art. 35); regulatory oversight of their algorithms and to provide public interest researchers with access to data to enable independent scrutiny of platform effects (Art. 40). Cf. The relevant provisions of the proposed AI Act, supra note 138, concerning high-risk AI systems. For a critique of the democratizing potential of such provisions, see however Rachel Griffin, \textit{Rethinking rights in social media governance: human rights, ideology and inequality}, 2 EUROPEAN LAW OPEN 30, 37-51 (2023).
\item \textsuperscript{155} Cf. Rachel Griffin, \textit{Public and private power in social media governance: multistakeholderism, the rule of law and democratic accountability}, 14 TRANSNAT’L LEGAL THEORY 46, 77-89 (2023); Hannah Bloch-Wehba, Algorithmic Governance from the Bottom Up, 48 BYU L. Rev. 69 (2022). For the contestatory role of digital activists and NGOs and their contribution to the effective constitutionalization of norms in digital spaces, see Christoph Graber, Net Neutrality: A Fundamental Right in the Digital Constitution?, in this symposium issue. For a co-design proposal concerning digital currency, based on societal constitutionalism and LPE scholarship, see Roxana Vatanparast, \textit{Digital Monetary Constitutionalism: The Democratic Potential of Monetary Pluralism and Polycentric Governance}, in this symposium issue.
\item \textsuperscript{156} Cf. Wong & Floridi, supra note 2.
\end{itemize}
limits of this approach are many and evident.

First of all, the implementation of digital technologies aimed at reducing the mental health distress of young users—especially the most vulnerable ones—runs contrary to the business model of most social media companies, with significant impacts on their profitability.\(^{157}\)

Importantly, this would affect not only Meta’s attractiveness for investors but also the broader economic ecosystem gravitating around the data analytics, high-level engagement, and virality allowed by social media networks. For example, one might think of the economy of social media influencing.\(^{158}\)

This element is not necessarily good news for the constitutionalization of digital platforms. As Kate Klonick stressed, the decline of profitability reduces their economic power but leaves their relational power largely unaffected.\(^{159}\) Most importantly, such decline might make them less sensitive to external pressures and less prone to implement social planner initiatives.\(^{160}\)

Actually, faced with the multiple setbacks to its ad-based business model,\(^{161}\) Meta is going in a direction opposite to that suggested here, as it is speeding up the launch of AI-driven advertisement.\(^{162}\)

Besides these considerations, taking the OB “too seriously” is a risk...
in itself, to the extent strategic actors (and scholars) could be enchanted by the “siren call of content moderation formalism,” which does not capture the scale and complexities of the digital spaces and leaves problems of legitimacy largely unresolved. In this sense, shaping litigation strategies before the OB is by no means intended to minimize the absolute necessity for public, politically legitimate regulation—at both the national and international level—as long as the latter takes into account the specificities of the normativity emerging from the digital space.

Against this background, however, and while preserving some healthy skepticism, it is also important to stress that scholars—especially engaged scholars—play a role in laying the analytical and normative preconditions for effective constitutional solutions. Through legal discourse, they provide a critical pool of argumentative solutions which might trigger or contribute to desired sociolegal evolution. In this sense, individuating and exploiting the transformative potential of an arena such as the OB to raise awareness on pressing issues, thematize global constitutional questions, and trigger conflict-driven rulemaking is not incompatible with other approaches. This applies particularly to right-based strategies. Indeed, although mostly thought to protect individual claims against states and governmental power, subjective rights, under certain conditions, might well unfold their trans-subjective potential to help address systemic issues. Perhaps, this article, like the entire symposium issue, could contribute to this endeavor.


165. See again Douek, supra note 48; and, more recently, Rachel Griffin, *Rethinking rights in social media governance: human rights, ideology and inequality*, 2 *EUROPEAN LAW OPEN* 30 (2023). See also, beyond the field of digital constitutionalism, CHRISTOPH MENKE, *CRITIQUE OF RIGHTS* (2020). One can also note that the question of the limits of rights-based litigation before judicial review bodies as a way to address systemic issues is a general problem of so-called transformative constitutionalism, also in state legal systems. See recently Armin von Bogdandy & René Urueña, *International Transformative Constitutionalism in Latin America*, 114 Am J. Int’l L. 403 (2020).

The Cost of Looking Good: How Fashion and Trend-based Consumerism Impact the Economy, Law, and Environment

ALLISON DENTON

I. INTRODUCTION

The apparel industry, and the importance of clothing, has transformed over time—whether clothing is used to protect the body from the elements or to express individuality, it is inescapable. Although fashion and the concern over appearances may seem frivolous to some, the industry itself is not immaterial. The global fashion industry has a value of almost $3 trillion, which is 2 percent of the world’s GDP.1 In 2017, consumers in the United States alone spent over $300 billion on clothing.2 The COVID-19 pandemic resulted in increased online spending, and the prevalence of e-commerce is as big as ever. Predictions that US consumers will break records in 2022 by surpassing $1 trillion in online spending3 seemed accurate, as Black Friday shoppers spent over $9 billion in one day.4 In February 2023, reports showed that 2022 ecommerce sales did break records by totaling $1.03 trillion.5 The size and prevalence of the fashion industry guarantees its impact on the global economy, the environment, and regulatory schemes; however,

unfortunately, much of the industry’s impact is negative.

A main contributor to the massive apparel industry is the fast fashion market. This sector of the industry caters to large audiences, many of whom are trend-focused consumers who want fashionable, inexpensive clothing at rapid speeds. Consequently, this segment of fashion promotes overconsumption by encouraging habitual purchasing behavior and perpetuating quick disposal of unpopular styles. The fast fashion business model creates detrimental effects that, without adjustment, will have catastrophic, everlasting impacts.

This note argues that implementing stricter regulations in the fast fashion industry would help mitigate its negative impacts. Although some countries have implemented minor restrictions that encourage fashion brands to better manage their waste production, more stringent action is necessary to protect the environment. Additionally, the fashion industry has broadly received backlash for its treatment of workers, and many fast fashion brands remain major perpetrators of using offshore, inexpensive labor. Moreover, the lackluster protection for luxury brands and fashion houses in the intellectual property field allows fast fashion brands to profit on technicalities and loopholes in the present law.

Section II will provide a historical overview of fashion and consumption. Subsection A will examine how significant fashion has been throughout time while analyzing the way clothing is utilized and its impact on society and individuals. Subsection B will focus on the shift in fashion consumption, investigating modern consumption and the transformation of consumer purchasing behavior, including trend-based consumerism.

Section III will delve into how the fast fashion industry operates. Subsection A covers the power and influence of social media. The popularity of social media platforms, namely YouTube, Instagram, and TikTok, has bolstered trend-based consumerism and regularly promotes fast fashion brands. Subsection B investigates the inner workings of the fast fashion industry, including the economic impact of the industry and its workforce, the intellectual property issues that are prevalent, and the environmental effects of the fast fashion model.

Finally, Section IV analyzes efforts to counteract the industry’s burden and includes proposals for new regulations and plans to mitigate its impact. Subsection A studies consumer awareness and purchasing behavior movements. It also looks at brand campaigns that promote sustainability. Subsection B examines various countries’ legislative action to counteract fast fashion’s impact and proposes new taxes and regulations to aid in this objective.
A. What is Fashion?

Articulating the meaning of fashion is difficult due to its various usages and meanings; however, Merriam-Webster defines “fashion” as a prevailing custom or style during a particular time. While correct, the concept of fashion is much more than just a popular style or a simple definition. It is a method of communicating ideas, dreams, and interests through clothes. Fashion has been coined “the cultural construction of the embodied identity,” as it allows individuals the opportunity to express who they are, who they want to be, and where they want to belong. The intimate relationship between fashion and an individual’s identity dates back thousands of years, long subjecting it to gossip, social stigma, laws, and royal decrees.

Research has shown that humans wore clothing as far back as 170,000 years ago. However, for many of those years, when civilizations were still primitive, clothing was typically worn for protection and concealment. Once humans began settling into well-defined social organizations— with art, culture, and languages—the ideology surrounding fashion began to change. This shift was prevalent in the Ancient World, including societies in Mesopotamia, Egypt, Greece, and Rome. At this time, styles differed by cultures and changes occurred slowly over hundreds of years. Similar styles remained prevalent throughout the Middle Ages, but by the early years of the Renaissance,
beginning in the fifteenth century, fashion began changing at a faster pace.\textsuperscript{16}

The emergence of fast-paced fashion resulted from economic mobility and new technologies—namely, modern tailoring.\textsuperscript{17} Prior to modern tailoring, clothing was made in the home, not purchased; however, once people began wanting certain styles of clothing, there was a demand for skilled tailors.\textsuperscript{18} Some historians have deemed the development of modern tailoring as the “birth of fashion.”\textsuperscript{19}

Following this development, clothing took on a new meaning, leading to new social implications and, later, legal restrictions. King Henry VIII passed “An Act Agaynst Wearing of Costly Apparrell,” which restricted clothing of “prestigious colors, refined quality, and exotic place of manufacture” to those of high status or nobility.\textsuperscript{20} It went further by assigning various colors, cloths, and designs based on ranking or class.\textsuperscript{21} Queen Elizabeth I implemented similar laws, limiting the type and amount of fabric available to different classes.\textsuperscript{22} These laws were hard to enforce, but violators were subject to fines and private individuals were able to “seize any apparel worn contrary to the statute . . . .”\textsuperscript{23} Here, fashion was being regulated to preserve the image and status of royalty and nobility. Along with the Crown, the Church, in its effort to defend traditional values, expressed opinions about fashion’s negative influence in promoting temptations like sex and vanity.\textsuperscript{24} Regulations of this kind were a strategic power play by institutions in an effort to perpetuate the social framework that had dominated for centuries.

Centuries later, during colonial expansion in the eighteenth century and the Industrial Revolution in the nineteenth century, new developments and technologies changed fashion once again.\textsuperscript{25} Advancements in production made clothing available to mass markets and, with a new class of buyers who could afford extravagant garments, the accessibility to fashion expanded.\textsuperscript{26} The shift to making clothes for mass markets quickly led to a standard of attire for certain social groups

\textsuperscript{16} Ford, supra note 7, at 38.
\textsuperscript{17} Id. at 39 (outlining the development of modern tailoring in the fourteenth century); see also The History of Tailoring, ALABAMA CHANIN JOURNAL (May 5, 2016), https://journal.alabamachanin.com/2016/05/the-history-of-tailoring/.
\textsuperscript{18} The History of Tailoring, supra note 16.
\textsuperscript{19} FORD, supra note 7, at 40.
\textsuperscript{20} Id. at 33.
\textsuperscript{21} Id.
\textsuperscript{22} Id. at 34.
\textsuperscript{23} Id.
\textsuperscript{24} Id. at 47.
\textsuperscript{25} See id. at 101.
\textsuperscript{26} Id.
and events. To ensure that class and status were still recognizable, dress codes developed detailing which outfits were appropriate for early mornings, visiting with acquaintances, afternoon promenades, evening dinners, and fancy balls.\textsuperscript{27} After time, however, the distinctions became less noticeable, shifting to a more streamlined and functional form of fashion,\textsuperscript{28} leading to the menswear staple, the classic black suit.\textsuperscript{29} Although menswear became simplified, this shift did not occur in womenswear for many more years.\textsuperscript{30}

There was no mass production of womenswear until after the Civil War,\textsuperscript{31} but even so it took decades before any real reform of women’s clothing occurred.\textsuperscript{32} The women of the early twentieth century swapped layers of material for lightweight fashions that skimmed the body.\textsuperscript{33} The new wave of fashion for women brought about ideals of personal freedom, sexual liberation, and recreation—a concerning shift for those attached to conventional gender roles and standards.\textsuperscript{34} Unfortunately, society was quick to disparage the flapper initiative, going as far as to suggest abolishing “all section of criminal law pertaining to rape” if women insisted on wearing the new provocative styles in public.\textsuperscript{35} Amidst the backlash, the change in womenswear forged ahead, continuing to evolve in time with feminist movements and women’s liberation.\textsuperscript{36}

World Wars I and II provided a major catalyst for the transformation of womenswear. During the wars, the men were sent off to fight and the women had to take on jobs typically occupied by men,\textsuperscript{37} like working in munitions factories; performing administrative work; and serving as drivers, nurses, and farm hands.\textsuperscript{38} Entering the workforce led to a reform of womenswear—promoting functionality and comfort.\textsuperscript{39} Subsequently, by the end of World War II, it became more common and acceptable for

\begin{itemize}
\item \textsuperscript{27}Id. at 104.
\item \textsuperscript{28}See Ford, supra note 7, at 117.
\item \textsuperscript{29}See id. at 102.
\item \textsuperscript{30}Id. at 137.
\item \textsuperscript{31}Id.
\item \textsuperscript{32}See id. at 143.
\item \textsuperscript{33}See id. at 143–44.
\item \textsuperscript{34}Id. at 146.
\item \textsuperscript{35}Id. at 148.
\item \textsuperscript{36}See id. at 152–53.
\item \textsuperscript{38}Id.
\end{itemize}
women to wear slacks.\textsuperscript{40} However, the long coveted privilege of being members of the workforce was soon clouded by misogynistic workplace behavior and standards. In the light of the recent Me Too movement, it is evident there are pervasive issues involving sexual harassment in the workplace and the standard set for employees.\textsuperscript{41} Another issue among workers is the discriminatory application of dress codes.\textsuperscript{42} Typically dress codes at workplaces or schools are unfairly enforced, often promoting different standards for men and women. In fact, evidence shows that women and girls are reprimanded and disciplined at disproportionate rates.\textsuperscript{43} This trend proves that, even now, after centuries of progression, fashion and womenswear continue to be regulated.

In addition to women, minorities also experienced regulations governing their attire.\textsuperscript{44} For example, laws were put into place, like South Carolina’s Negro Act of 1740, that restricted slaves from wearing or owning anything of higher quality or value than of the “condition of a slave.”\textsuperscript{45} These restrictions created a racial dress code, further perpetuating the social status and class of certain groups. Once slaves were emancipated on June 19, 1865, many freed slaves “toss[ed] their ragged garments into the creeks” as a way to cast off their former identity as a slave.\textsuperscript{46} Decades later during the Civil Rights era, the Black community used fashion as a means to communicate its message.\textsuperscript{47} For marches, demonstrations, and protests, the message of respectability was

\begin{thebibliography}{99}
\footnotesize
\item \textsuperscript{40} Id.
\item \textsuperscript{44} Ford, \textit{supra} note 7, at 157.
\item \textsuperscript{45} Id. (quoting Shane White & Graham White, Slave Clothing and African-American Culture in the Eighteenth and Nineteenth Centuries, 148 PAST & PRESENT 149, 159 (1995)).
\end{thebibliography}
at the forefront, as it was “intended to elevate the Black community in the eyes of the greater public.” To convey this message, activists wore their Sunday best, forcing people to “move beyond the stereotypes of blacks and black bodies.” The clothing worn and the appearances presented during this pivotal time had profound meaning in the political realm. Similarly, fashion and aesthetics played a prominent role in the Black Power agenda, with groups like the Black Panther Party adopting a distinct uniform. The Civil Rights and Black Power movements strategically used fashion, knowing its power to communicate messages in “the battle to shift American public consciousness on matters of race.”

Fashion has not lost its weight over the years, and now with the growth of social media and worldwide connection, fashion is as important as ever since it reaches a far wider group of people at a much quicker rate. For thousands of years, fashion has been a pervasive and influential tool for communication. Whether people dress for mere utility or utilize fashion as a device for self-expression, the fashion industry is here to stay. Instead of trends taking hundreds of years to progress, trend cycles occur almost every twenty years. However, with the upsurge in microtrends, which sees certain items gain popularity at astronomical rates and fall into disfavor just as quickly, the fast fashion market has thrived—resulting in devastating consequences. The fashion industry will endure, and changes must be made to counteract its negative implications.

B. An Overview of Fashion Consumption

As the fashion industry changed, so did the approach to how people consume clothing. According to reports, every US household spends an average of $1,700 on clothes per year. Additionally, the global fashion

48. Id.
50. Id. at 178–79.
51. See id. at 183; Donaldson, supra note 47.
52. Donaldson, supra note 47.
industry amounts to almost $3 trillion, which is two percent of the world’s GDP.\footnote{Vilaça, supra note 1.} Many countries in the world participate in this industry, but just ten dominate the market.\footnote{Volume and Consumption: How Much Does the World Buy?, COMMON OBJECTIVE (May 14, 2018), https://www.commonobjective.co/article/volume-and-consumption-how-much-does-the-world-buy.} China, the United States, India, Japan, Germany, the United Kingdom, Russia, France, Italy, and Brazil account for three-quarters of the clothes sold each year.\footnote{Id.} This massive industry has continued to grow over the past decades and is predicted to continue to rise. Analyzing the shift of consumerism over time puts these numbers into context and highlights the enormity of the industry.

First, looking at the eighteenth century, the rate of consumption and the cost of attire was drastically different than modern times. Marie Antoinette, the last queen of France before the French Revolution, is famous for her luxurious lifestyle and indulgent spending habits.\footnote{Chloe Foussaines, Why Marie Antoinette Still Reigns Over Fashion, TOWN & COUNTRY (Feb. 28, 2020), https://www.townandcountrymag.com/style/fashion-trends/a31102612/marie-antoinette-fashion-influence-explained/.} She had a much larger budget than a commoner’s. Comparing budgets gives insight into consumer patterns. In today’s dollars, Marie Antoinette’s clothing allowance was $3.6 million, but she often doubled this amount, equating to roughly $20,000 a day on clothing.\footnote{Carrie Kirby, How Much Would It Cost to Actually Live Like a Queen for a Day?, WISEBREAD (May 26, 2016), https://www.wisebread.com/how-much-would-it-cost-to-actually-live-like-a-queen-for-a-day.} This amount is substantial and excessive; comparatively, nonroyals bought clothing differently, and, much like today, their budget dictated the size of their wardrobes.\footnote{Loretta Chase & Susan Holloway Scott, How Many Clothes Did an 18th C. Woman Own?, TWO NERDY HISTORY GIRLS (Sept. 16, 2014), https://twonerdyhistorygirls.blogspot.com/2014/09/how-many-clothes-did-18th-c-woman-own.html.} Those who would be appearing at court or attending balls typically still needed to wear the appropriate outfits in the latest styles. It has been estimated that one court dress could cost a minimum £200 and could go up to £500 or more.\footnote{Regina Jeffers, Costs of Living During the Regency Period, EVERY WOMAN DREAMS (Feb. 24, 2017), https://reginajeffers.blog/2017/02/24/cost-of-items-during-the-regency-period/ (calculating to above $12,000 in today’s money).} However, for the majority of people, before the twentieth century, they could only afford one new outfit a year, and “the poor sometimes had to go without underwear or pawn their winter clothing in summer.”\footnote{How Much Did a Shirt Really Cost in the Middle Ages?, BOOK AND SWORD (Dec. 9, 2017), https://www.bookandsword.com/2017/12/09/how-much-did-a-shirt-really-cost-in-the-middle-ages/.}
worth of income on a single item.\textsuperscript{64}

Once mass production made clothes more accessible, the way people bought apparel transformed. The mass production of clothing increased and it became an established industry a few years into the twentieth century.\textsuperscript{65} By 1915, the apparel business had grown rapidly, even ranking as the third largest American industry, following the steel and oil industries.\textsuperscript{66} The impact of the Industrial Revolution and the growth of the apparel industry are evident as they allowed “even people of average means to have a number of outfits for different occasions, and continue to buy clothes simply because of their tastes.”\textsuperscript{67}

Although clothing was more accessible, it was still not purchased at the same rate it is now. In 1930, the average American woman owned nine different outfits.\textsuperscript{68} By the 1960s, the average American household spent over ten percent of its income on apparel and footwear, which is roughly $4,000 today, and the average person bought fewer than twenty-five garments each year.\textsuperscript{69} At that time, almost all clothing was made in the United States, around 95 percent,\textsuperscript{70} compared to 2013 in which only around 2 percent of clothing was produced in the United States.\textsuperscript{71}

Furthermore, as mentioned, the average household now only spends $1,700 on clothes and shoes per year, less than 3.5 percent of its budget. Although we spend less money proportionally, because of a decrease in price and an increase in convenience, nearly twenty billion garments are purchased a year.\textsuperscript{72} This equates to around seventy items per person, or more than one fashion purchase every week.\textsuperscript{73} Additionally, the quality of clothing has diminished in the last few decades, which means garments deteriorate at a much quicker rate, leading consumers to replenish their closets more often.\textsuperscript{74}

The increase in consumption can in part be contributed to the rise of

\textsuperscript{64} Id.
\textsuperscript{66} Id.
\textsuperscript{68} Id.
\textsuperscript{70} Id.
\textsuperscript{71} Id.
\textsuperscript{72} Id.
\textsuperscript{73} Id.
the fast fashion industry, with popular brands like H&M, Zara, Shein, and Fashion Nova dominating the market. Fast fashion consists of cheap, trendy clothing that “samples ideas from the catwalk or celebrity culture and turns them into garments in high street stores at breakneck speed to meet consumer demand.” This business model allows consumers to buy fashionable items conveniently, wear them while in style, and discard them after a few wears. Another contributing factor is “trend-based consumerism,” coined for the purposes of this note, which encompasses the idea of how the modern consumer interacts with the fashion industry. Following trends has been a foundation of the fashion industry for many years, but with the growth of social media and online shopping it has taken on a new form. The rapid rotation of trends and microtrends spotlights certain products—leading to consumers purchasing and often selling out items in as little as a day. This new shift in consumerism and fashion consumption is unlike previous conduct and as such must be analyzed in relation to the prosperity of the fast fashion industry to which it strengthens.

The fast fashion model and the significant increase in consumption is attributed to various factors—outsourced labor, low quality materials, social media and online retailers, etc.—which will be discussed in depth throughout this paper.

III. THE POWER OF FAST FASHION

A. The Influence of Social Media and Trend-Based Consumerism

Social media plays a large role in the modern fashion industry and, specifically, the success of the fast fashion sector. Social media slowly gained traction over several years before it skyrocketed into popularity.  

75. “High street stores” is a common UK term that refers to the clothes you could buy on your average UK street, including brands that mass produce clothing that caters to large audiences and are affordable. This is comparable to the term “fast fashion” which will be used in this note. See Vaishnavi Thakur, Do You Know the Difference Between High Street and Couture Brands, BE BEAUTIFUL (Oct. 23, 2017), https://www.bebeautiful.in/fashion/trends/difference-between-high-street-and-couture-brands; Solene Rauturier, What is Fast Fashion?, GOOD ON YOU (July 26, 2021), https://goodonyou.eco/what-is-fast-fashion/.

76. Rauturier, supra note 75.

77. See generally Seren Morris, 20 Products that have Sold Out Because of TikTok Popularity, NEWSWEEK (Apr. 9, 2021), https://www.newsweek.com/20-products-sold-out-because-tiktok-popularity-1581946 (showing the impact of TikTok on the market).

78. See generally, History of Social Media (It’s Younger Than You Think), BROADBAND SEARCH, https://www.broadbandssearch.net/blog/complete-history-social-media (last visited Jan. 28, 2023) (giving a historical overview of social media).
but nowadays, social networks, online communities, blogs, and forums have become a part of daily life for many people. Studies have shown that consumers gather fashion ideas from media, movies, social networks, and friends. Sharing and connecting with other consumers online about these ideas quickly turns them into trends, which is one reason social media is viewed as the key to reach large audiences, develop brand awareness, and popularize items. Additionally, the increased presence of social media provides brands a unique platform to connect with consumers and influence purchasing decisions. For example, on Instagram, popular fast fashion brands Zara and H&M have fifty-five million and thirty-eight million followers, respectively. The immense influence of social media is apparent, and social analysis is important “when measuring consumer conversation and interactions on Facebook, Instagram, and other social media channels.”

Brands capitalize on the popularity of social media by investing in digital marketing strategies. The use of digital media marketing has become increasingly prevalent as social media and e-commerce have expanded. Digital marketing is comprised of various channels for advertisers to utilize when communicating with consumers. The main channels are social media marketing, influencer marketing, affiliate marketing, content marketing, display advertising, email marketing, search engine marketing, video advertising, and mobile marketing. For the purposes of this note, social media and influencer marketing will be analyzed further in connection with the fast fashion industry.

Social media marketing involves using social media channels, such as Instagram, Twitter, YouTube, TikTok, and Facebook, to promote products. This strategy “draw[s] in traffic for a website or . . . gain[s] attention of buyers over the web using different social media platforms.” The ability of fashion brands to advertise on social media provides ads of popular items to consumers at the tap of a button. It also expands a brand’s market, allowing international brands to find footing in new markets.

80. Id.
82. Michaela, supra note 79, at 174.
84. See id.
85. Id.
markets.\textsuperscript{86} Social media marketing has various methods that brands can adopt to best fit their goals. One method is to capitalize on the amassed following of popular influencers.

Influencer marketing is a creative approach to combining the concept of celebrity endorsements and modern-day content creators.\textsuperscript{87} After identifying an influencer, brands will engage with those individuals “who have an influence over a specific target audience or medium, in order to be part of a brand’s campaign towards increased reach, sales, or engagement.”\textsuperscript{88} Influencer marketing has been consistently prevalent throughout the last decade, gaining recognition at the height of fame for YouTube creators, and now partnering with Instagram and TikTok creators. The use of influencer marketing can be divided into two subparts: earned (or unpaid) and paid. Earned marketing allows brands to profit from creators talking about the brand naturally and organically, growing positive brand recognition.\textsuperscript{89} Paid marketing campaigns involve compensating influencers to talk about the brand or products, with brands using this form of marketing to reach their goals, “establish credibility in the market, . . . create social conversations around their brand . . . [or] drive online or in-store sales of their products.”\textsuperscript{90}

Influencers have a lot of weight when it comes to their followers’ purchasing decisions, and the FTC implemented a law to regulate influencers’ online endorsements. Under the FTC’s law, content creators have a legal obligation to “clearly and conspicuously disclose their material connection to the products, services, brands, and/or companies.”\textsuperscript{91} The FTC’s regulation of social media advertisements has made consumers aware of an influencer’s motivations for sharing particular products or supporting certain brands.

Many marketing options are beneficial for companies, but influencer marketing has become the top choice for many brands. On average, brands generate $6.50 in revenue for each dollar invested in influencer marketing, with 13 percent of brands making twenty dollars

\textsuperscript{86} See id.
\textsuperscript{88} Sudha, supra note 83, at 16.
\textsuperscript{90} Sudha, supra note 83, at 16.
This economic benefit has incentivized brands to create influencer marketing campaigns and use all available platforms. One prominent platform is TikTok, a fast-paced video sharing network, with around one billion users in more than two hundred countries, and around 90 percent of users accessing the app daily. On a global scale, this app is massive, and it has a large presence among younger users. Thirty-five percent of TikTok’s global audience is between the ages of nineteen and twenty-nine, and 28 percent is under eighteen. Many of those users fall into Gen Z, those born from 1997 to 2012, making them the largest generation of trendsetters, especially in fashion.

Brands have used TikTok creators’ large audiences to push their products, but another phenomenon has taken over the app. The surge in fashion haul videos, where a user shares all of his or her recently purchased items, has reinforced trend-based consumerism. “Hauls are a subtle, yet entertaining way of implementing consumerist culture and pushing the idea of a consumerist society onto an audience.” These hauls encourage consumers, and especially content creators, to frequently buy popular items and share it with their followers. Shein, an international fast fashion e-commerce brand, is showcased in many of these haul videos trending on TikTok and YouTube. In fact, a quick search of #sheinhaul on TikTok boasts videos that collectively have over 2.6 billion views. Due to the low cost of the garments and the vast array of trendy options, some haul videos feature up to $800 worth of clothes, and, at the extreme, some videos can range into thousands of dollars. Not only do the haul videos glamorize overconsumption and promote shopping unsustainably, but also they advance the fast fashion model. A poll reported that 66 percent of people aged fourteen to twenty-seven admitted to having

98. Id.
99. See id.
bought more clothing since watching TikTok. Notably, recent online shopping statistics revealed that American consumers use their mobile devices most often to buy items. And overall, mobile purchasing dominated 73 percent of all e-commerce sales in 2021.

With TikTok’s fast-paced design, trends rotate through quickly, and it “creates the perfect space for fast fashion brands to swoop in and cater to the speed of the trends.” Fast fashion brands, like Zara, have a product turnaround of two to six weeks, which allow them to accommodate consumers’ fluctuating interests. A report found that the brand Missguided releases around one thousand new pieces monthly, and Fashion Nova launches between six hundred to nine hundred new styles weekly. Famous designers and luxury brands release collections roughly two times per year, which compete with fast fashion brands’ fifty-two “microseasons” a year, figuring about one new “collection” a week. How exactly does the fast fashion industry operate at such a high-paced rate? By exploiting workers, taking advantage of lackluster intellectual property protection, and exhausting the environment.

B. The Foundation of the Fast Fashion Industry

1. Economic and Labor Implications

As discussed, the fashion industry is massive, and its effects are far reaching. The multibillion-dollar business profits off fast turnaround and low prices, but the cost of this business model is carried elsewhere. There are societal, human, and environmental costs that result from an industry structure solely focused on profit. These detrimental consequences, if left unchecked, will continue to worsen. A report found that addressing the environmental and social issues created by the fashion industry would “provide a $192 billion overall benefit to the global economy by 2030.”

100. Meehan, supra note 96.
102. Id.
103. Meehan, supra note 96.
Fashion is one of the most labor-dependent businesses, and it has been reported that one in six of the world’s workers are employed in the fashion industry. However, the fashion industry at large outsources labor and clothing production to maintain low prices. The fashion supply chain has changed dramatically over the last fifty years, with many design companies “extend[ing] their supply chains and us[ing] subcontractors to complete unskilled labor tasks.” Unfortunately, this often leads to unethical labor practices leaving workers subject to long hours without compensation and unsafe working conditions. Nevertheless, in order for brands to meet consumer demands and keep up with fast-paced trend cycles, companies are increasingly dependent on low-cost labor. Over the past few decades, the fashion industry has made the choice to outsource labor to low-income Asian countries, like Bangladesh, China, Vietnam, India, and the Philippines, where mainly female factory workers are recruited. In these countries, women are socially and economically vulnerable, allowing brands to take advantage of their need for employment.

This production structure leads to accountability issues as many companies use subcontractors to complete work internationally and leave workers and brands without legal recourse. This structure survives due to two main factors: low liability for fast fashion brands and consumer apathy. First, the lack of liability fast fashion brands are exposed to enables poor working conditions and unethical treatment of workers to perpetuate. Typically, design companies are only legally responsible for suppliers if the design company directly employs them. But, many companies contract with outside manufacturers, or middlemen companies, which is not enough to impose legal liability on the design company. This arrangement allows for poor working conditions in many offshore factories. Additionally, many of these manufacturing subcontractors “have no assets besides thread and needles, and they

108. See Nguyen, supra note 104.
110. Id.
111. Id.
112. Id.
113. Id.
114. Sutor, supra note 109, at 860.
115. Id.
116. See McCosker, supra note 107.
declare bankruptcy when employees sue.”

Horror stories coming out of these factories highlight the hardships workers, especially women, face. In Bangladesh, women reported that they developed bladder infections due to the lack of bathroom breaks, and some managers were even forcing workers to take contraceptive pills. This horrible treatment is exacerbated by the denial of maternity leave, extremely low wages, poor working conditions, and sexual harassment in the workplace. Women in these communities cannot afford to risk the loss of income, so they are increasingly more susceptible to sexual abuse or harassment and do not report it. Sadly, one in four Bangladeshi garment workers disclosed some form of abuse in a report to Oxfam. The workers need their income, but, according to a 2019 report by Oxfam, 0 percent of Bangladeshi garment workers and 1 percent of Vietnamese garment workers earned living wages. Due to inadequate wages, women often have to bring their young daughters—as young as ten—to work with them in order to feed their families. This system is extremely disadvantageous to women and although many consumers cherish the low cost of trendy items, those inexpensive garments are produced on the backs of poor, overburdened workers.

Additionally, workers are often exposed to harmful chemicals, which are linked to several kinds of cancer and respiratory issues. Workers’ health and safety are constantly at risk when they are forced to work in unsafe buildings, are exposed to heat, are susceptible to dangerous chemicals, have no access to clean drinking water, and have restricted access to the bathroom. Factories have shown that there is little to no concern for the safety of their workers. For example, countless factories “have no emergency exits or they lock the doors and bar the windows to deter theft, causing many people to die when a factory fire occurs.” In 2012, a fire at the Ali Enterprises in Pakistan left workers trapped in the burning building and resulted in 254 people dying and over fifty more seriously injured. Similarly, in 2020, a fire broke out in an Indian factory, which manufactures clothing for Target, H&M, and Wal-Mart, and, because the factory had only one exit, multiple workers died as a result of being trapped.

117. Sutor, supra note 109, at 861.
118. See McCosker, supra note 107.
119. Id.
120. Id.
121. Id.
123. Id.
124. Id.
125. Sutor, supra note 109, at 879.
Association and various factory workers claimed that Indian factories rarely get inspected and few enforce safety regulations.\textsuperscript{126} Unfortunately, many brands are aware of these conditions and still choose to contract with these manufacturers because cost-benefit analysis proves that low-cost production and unethical business practices still result in massive sales, so brands feel dissuaded from making changes.

Secondly, consumers seem largely unaffected by the burdens of the fast fashion industry. In fact, it is rare that fashion retailers lose large portions of their customer base due to poor labor practices.\textsuperscript{127} Public criticism can spur change and serve as a catalyst for improvement, but many consumers do not seem to keep in mind a brand’s unethical practices when purchasing from them.\textsuperscript{128} Also, people tend to prioritize convenient purchasing and affordable pricing over sustainability.\textsuperscript{129} Notably, even when brands come under fire for labor malpractices, it creates almost no significant shockwaves—big celebrities still endorse the brands and customers still purchase clothing from them.\textsuperscript{130}

The fast fashion industry must make improvements in its labor practices, and even minor changes to the structure of production can have exceptional outcomes for workers. Deloitte Access Economics found that paying fashion workers a living wage throughout the supply chain would likely only increase the retail price of a garment by 1 percent.\textsuperscript{131} Another study, conducted by researchers Hall and Wiedmann, found that increasing the cost of clothing pieces made in India an average of twenty cents per item would be enough to raise all Indian garment workers out of poverty.\textsuperscript{132} These numbers illustrate the reasonable measures that can create remarkable changes.

\section*{2. Intellectual Property Issues}

Intellectual property (IP) law is an area of law that deals with protecting and enforcing the rights surrounding inventions, writing, music, designs, and various other works.\textsuperscript{133} Several specific areas make

\begin{itemize}
\item \textsuperscript{126} Id. at 880.
\item \textsuperscript{127} Nguyen, supra note 104.
\item \textsuperscript{128} Id.\textsuperscript{129} Id. This finding is based on a 2018 report that surveyed nearly 700 shoppers ages 18 to 37.
\item \textsuperscript{130} Id.\textsuperscript{131} Id.\textsuperscript{132} McCosker, supra note 107.
up IP law, including copyright, trademark, and patents.\textsuperscript{134} However, this field of law is an imperfect avenue for fashion designers to protect their designs.\textsuperscript{135}

Copyright law protects “original works of authorship,” including books, paintings, sculptures, and songs, but it does not protect useful things, like clothing and accessories.\textsuperscript{136} Due to this limitation, it provides little protection for apparel in its entirety. However, some creative elements of a design may be separated from the functional aspect of the garment,\textsuperscript{137} such as the prints of fabrics, and are afforded protection under copyright law.\textsuperscript{138} Obviously this narrow protection does little to combat the plethora of issues in the fast fashion industry, but other sectors of IP law offer a little more protection to the apparel industry.

Namely, trademark law offers more, but not great, protection for fashion designers. Trademark law safeguards “any symbol, work, color, or other feature that identifies a good in commerce.”\textsuperscript{139} However, trademark law requires the protected characteristic to “achieve a secondary meaning in trade” and fashion designers often fail to protect their creations under these rules.\textsuperscript{140} Trademark law links a brand’s name and logo to its products to further brand identity and reinforce its reputation for quality items. Therefore, a main part of trademark law serves to protect the exclusivity and high status of certain items.\textsuperscript{141} Even when items are clearly marketed as designer knockoffs, and most consumers are likely aware of what they are buying, to casual third parties these items may appear real.\textsuperscript{142} This dilutes the market for luxury, high-status items and makes actual expensive items less unique.\textsuperscript{143} For example, Christian Louboutin, a famous shoe designer known for the red soles of his shoes, is entitled to bar other shoes with red soles, even if they are branded and marketed differently.\textsuperscript{144} An influx of red-soled shoes would undermine the association of the Christian Louboutin brand with wealth and high status.\textsuperscript{145} Consumers typically

\begin{itemize}
\item \textsuperscript{134} Id.
\item \textsuperscript{135} Sutor, supra note 109, at 863.
\item \textsuperscript{137} Id.
\item \textsuperscript{138} Id.
\item \textsuperscript{139} Sutor, supra note 109, at 863.
\item \textsuperscript{140} Id. at 883.
\item \textsuperscript{141} See Ford, supra note 7, at 299.
\item \textsuperscript{142} Id.
\item \textsuperscript{143} Id.
\item \textsuperscript{144} Id.
\item \textsuperscript{145} Id.
\end{itemize}


149. INSTAGRAM, FashionNova, https://www.instagram.com/p/BmUXT9GhC-M/?utm_source=ig_embed&ig_rid=cca4c69e-2a30-416d-8d71-cf239f2d27f9.}
Since the fast fashion business model relies on the ability to quickly recreate high-end trends at an affordable price—and does so by copying other designers—it is necessary to implement stricter protection and legal consequences for fast fashion brands.150

3. Environmental Considerations

Obtaining affordable, accessible fashion is a luxury, and many consumers fail to consider what it takes to produce their garments. The fashion industry generates exponential water waste, harmful chemical exposure, and damage from the use of low-quality materials.

To put into perspective just how impactful the fashion industry is, note that approximately 20 percent of water waste worldwide is

150. Cohen, supra note 146.
attributable to the textile dyeing process and production system.\textsuperscript{151} Every year, the fashion industry alone uses over twenty-four trillion gallons of water—enough to meet the consumption needs of five million people.\textsuperscript{152} More specifically, it takes over seven hundred gallons of water to produce one cotton shirt.\textsuperscript{153} Strikingly, the amount of water it takes to make one shirt is enough to meet the average person’s drinking needs for two and a half years.\textsuperscript{154}

Additionally, fast fashion brands use cheap materials, many of which are toxic or are dyed using toxic materials. In fact, the fashion industry is the second largest (following the agricultural industry) polluter of clean water globally.\textsuperscript{155} Also, many brands commonly use cheap textiles, like polyester, which are derived from fossil fuels and adds to global warming by shedding microfibers that contribute to the increasing amounts of plastic in the oceans.\textsuperscript{156}

These microfibers are essentially microscopic pieces of plastic in garments that enter the water supply during their production cycle or after purchase by washing them. They are extremely damaging to the environment—half a million tons of microfibers are dumped into the ocean yearly, the equivalent of fifty billion plastic bottles.\textsuperscript{157} However, even textiles that are considered more “natural,” like cotton, require the use of enormous amounts of water and pesticides.\textsuperscript{158} Since most production is done in developing countries, this creates a competitive atmosphere where local communities have to compete with large companies for resources.\textsuperscript{159}

On top of water waste, the fashion industry is culpable for 10 percent of annual global carbon emissions, which is “more than all international flights and maritime shipping combined.”\textsuperscript{160} Unfortunately, if this rate continues, the fashion industry’s greenhouse gas emissions will increase

\textsuperscript{153} See Le, supra note 151.
\textsuperscript{154} Reichart & Drew, \textit{supra} note 106.
\textsuperscript{155} Rauturier, \textit{supra} note 75.
\textsuperscript{156} Id.
\textsuperscript{158} Rauturier, \textit{supra} note 75.
\textsuperscript{159} Id.
\textsuperscript{160} The World Bank, \textit{supra} note 152.
by more than 50 percent by 2030.\textsuperscript{161} For context, making one pair of jeans produces the same amount of greenhouse gases as driving more than eighty miles.\textsuperscript{162}

Furthermore, due to the increased rotation of trend cycles and the low quality of fast fashion garments, clothing is disposed of at a disastrous rate. Every year, Americans throw away approximately eighty pounds of clothing, occupying almost 5 percent of space in landfills.\textsuperscript{163} This discarded clothing, especially those made of nonbiodegradable materials, can end up sitting in landfills for up to two hundred years.\textsuperscript{164}

The senseless and irresponsible damage done by the fast fashion industry must be combatted with regulatory action. Although some countries have implemented preventative measures, more regulatory systems are needed to encourage improvements and enforce change.

IV. EFFORTS TO COUNTERACT THE FAST FASHION INDUSTRY

A. Increasing Consumer Awareness

Recently, the fast fashion industry has received criticism for its damaging effects on the world, which has led to movements aimed at increasing consumer awareness. The movements focus on exposing the fast fashion industry’s use of poor labor practices, highlighting its negative environmental output, and encouraging sustainable purchasing behaviors among consumers.

Although consumer interest and desire for more sustainable options has increased the market for ethically sourced clothing, more improvements are required. Popular fast fashion brands have responded to the increased interest in sustainability by coming out with “ecofriendly” options. But it is important to note that many of these brands use “greenwashing” techniques, which are frequently used marketing strategies designed to make products appear more sustainable than they actually are.\textsuperscript{165} Companies use these marketing tools to make consumers feel better about their purchases, while really not making any substantial changes on the production or labor front.\textsuperscript{166} Many brands are

\begin{itemize}
\item \textsuperscript{161} Id.
\item \textsuperscript{162} Reichart & Drew, supra note 106.
\item \textsuperscript{163} Rachel Bick et al., The Global Environmental Injustice of Fast Fashion. 17 ENV’T HEALTH 1, 1–2 (2018).
\item \textsuperscript{164} Reichart & Drew, supra note 106.
\end{itemize}
not that transparent with their sustainability progress. For example, a popular fast fashion brand, H&M, released a new “Conscious” collection which is advertised as a sustainable option because it is made of wool. In reality, the garment’s composition was only four percent wool. In October of 2022, the United Kingdom’s Competition and Markets Authority developed the Green Claims Code to counteract greenwashing. This Code enumerates six key points that evaluate if brands’ environmental claims are genuinely ecofriendly.

Additionally, some large brands offer “recycling” services that incentivize customers to donate or recycle clothes by offering discounts or loyalty points. Sadly, many of these programs do not follow proper reuse channels and the clothing either gets resold or dumped in low- or middle-income countries where unsold items are either burnt or left to sit in landfills. Using stricter regulatory processes and implementing transparency requirements for brands that, allegedly, sell “sustainable” options will increase brand liability.

Although regulatory actions and taxation will likely be the most effective solutions to bring about considerable change, consumers also play an important role. If consumers continue to support the fast fashion industry, the current system will continue to thrive, and companies will profit from billions of sales worldwide. Implementing various “nudges” to consumers will aide in shifting their purchasing behavior. A nudge is “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives.” This behavioral economics theory could surely influence consumer patterns. Placing sustainable clothing options at the front of stores, in window displays, and in advertisements will subtly expose consumers to new, ecofriendly—yet still accessible—options. Another useful tactic would be to take advantage of the vast power of social media by introducing marketing campaigns and influencer brand deals centered around sustainable fashion. This will encourage buyers to think about their purchases differently. Brands have to make the changes but, until customers stop supporting the current system and brands’ bottom lines are affected, no changes will be made.

Another option to nudge consumers to practice ethical behavior is to encourage donating and recycling clothing instead of contributing to the massive disposal rates. Although a tendency for shopping secondhand has

167. Id.
169. See id.; Bick, supra note 163, at 2.
become popular recently, with content creators showing thrift hauls or selling used clothing on resale sites, there is still a massive issue with clothing waste. Over the past few years, countries have implemented deposit return schemes (DRS) for drink containers, and mimicking this program structure to counteract fashion waste would be an excellent use of established recycling schemes. England, one among thirty-eight countries with these programs, introduced its DRS program, which was intended to decrease litter by incentivizing consumers to return their bottles and cans in exchange for a small cash sum. Many DRS programs use “reverse vending machines” to facilitate cataloging the items and giving consumers rewards. These programs have made a significant improvement in participating countries, increasing recycling rates to more than 90 percent. Using similar systems for recycling garments may help mitigate the vast textile waste issue that results from rapid trend cycles and wear-and-discard practices.

Implementing any of the above alternatives would provide opportunities for consumers to make better choices when shopping. By nudging consumer behavior, brand practices will be forced to change and adapt to a changing consumer focus.

B. Legislative Action

Due to the globalization of the fashion industry in the past few decades, it is much harder to have widespread control over the industry’s practices and standards. Because many fast fashion companies operate with the goal of earning maximum profits, offshore manufacturers are employed to cut costs. Countries where these manufacturers are based have little capability of implementing stringent regulations to combat pollution, hazardous waste, and labor mismanagement. Unfortunately, regulations in these countries are not feasible, in part because of their organizational structures and governments, but also because the changes would be largely ineffective. The employment markets in these countries thrive on the ability for fast fashion brands to cut manufacturing and labor costs, and increasing regulations would require costly restructuring and enforcement plans. Companies may also get around regulations

172. Id.
173. Id.
174. Id.
175. Ashley Lauren, Why Regulations Aren’t Solving the Fashion Industry’s Environmental Problem, MEDIUM (Nov. 11, 2019), https://medium.com/age-of-
from the United States or EU, for example, by continuing to shift their damaging behavior to developing countries that do not have the infrastructure to enforce ethical practices. These issues require a comprehensive global initiative to make changes in the countries where brands are headquartered and also in the countries where production and labor occur.

Some legislative action has been taken across the world to combat these large-scale issues. In 2017, the OECD, which currently consists of thirty-eight participating countries, adopted a “global guidance standard on companies’ due diligence requirements towards their supply chain as it relates to child labor, forced labor, workers’ wages and collective bargaining agreements.” This guidance, as its name suggests, is not legally binding, although OECD members are expected to abide by it. Additionally, the Sustainable Apparel Coalition, boasting 250 global members, provides a foundation for sustainable manufacturing.

Specific countries have begun introducing legislation to combat unfavorable behavior. In 2022, the Fashion Accountability and Building Real Institutional Change (FABRIC) Act was proposed in the United States in an effort to establish the United States as a global leader for ethical apparel production and labor practices. The Act’s aim is to amend the 1938 Fair Labor and Standards Act (FLSA) and forbid piece rate payment plans for garment industry employees. This payment structure is determined by how many garments a worker can produce, with each garment earning between two and six cents. This payment structure creates considerable obstacles for workers to earn a minimum wage. The FABRIC Act would establish a nationwide registry through the Department of Labor in an effort to promote “transparency, hold bad

awareness/why-regulations-arent-solving-the-fashion-industry-s-environmental-problem-9a50be4c2843.

176. OECD is the Organisation for Economic Co-operation and Development which is dedicated to creating better policies and improving quality of life. See generally OECD, https://www.oecd.org.


178. Lauren, supra note 175.


actors accountable, and level the playing field.” New requirements for fashion brands, retailers, and manufacturing partners would hold them jointly liable for workplace wage violations.

In a 2019 report, the Environmental Audit Committee (EAC) proposed retail industry-specific sustainability recommendations to the UK government. The EAC recommended that tax reforms could reward retailers for reducing their carbon footprint, but the UK government has not acted on any of the proposed options as of yet. It did acknowledge the need for change in the upcoming future. However, instead of reforming taxes to reward brands, implementing a Pigouvian tax may be more suitable. This type of tax imposes a tax “on a market transaction that creates a negative externality, or an additional cost, borne by individuals not directly involved in the transaction.” Initiating this type of tax will require expanding brands’ exposure to liability. It is inequitable to continue letting fast fashion brands profit by exploiting workers and avoiding liability by using middlemen manufacturers. Perhaps imposing legal liability on brands that conduct a majority of their business and production network by contracting with unethical offshore manufacturers will induce change.

Taxing big fashion brands for taking advantage of contracting with inexpensive middlemen is one option to improve the industry. It may also address the existing intellectual property law issues by disrupting the production channel. If fast fashion brands and independent manufacturers become responsible for improper working conditions, the cost of production will increase. In turn, it will decrease the profitability of copying designs since it will no longer be feasible to use unfair labor practices to rapidly produce them. According to WIPO Magazine, “the European Union has registered and unregistered Community design rights that provide protection for garments and accessories as a whole. That simply does not exist in the United States.” This type of tax would limit the availability of fast fashion brands use of IP loopholes.

Another avenue is to introduce taxes and bans on environmental waste. France appointed a Secretary of State to the Minister for the

---

183. *Id.*
184. Huckle, supra note 179.
185. *Id.*
Ecological and Inclusive Transition, who is dedicated to repairing the fashion industry and aiding in implementing antiwaste laws, which the French parliament passed in 2020. This piece of legislation introduced a ban on the burning of unsold clothing in France. Similarly, the Swedish government has approved instituting a tax on chemicals used in the fashion industry. A researcher at the Centre for Future Chemical Risk Assessment and Management Strategies noted that the proposed new tax would cover all garments and shoes and would allow for tax deductions, if companies could demonstrate the lack of harmful chemicals in their products.

Progress has been made in the fashion industry over the years, but it is minimal compared to the increase in sales and damaging effects. There is still much improvement to be done before any considerable sustainability goals will be reached. But, it is obvious that worldwide cooperation and dedication is a necessity if any efforts are going to prove fruitful.

V. Conclusion

The reach of the fast fashion industry is extensive and, without proper management, it will cause irrevocable harm. Fast fashion is supported by millions of consumers worldwide and global change must be made in order to see a shift in brand action. Implementing systemic changes, by using consumer nudge plans and regulatory action, will serve as a catalyst for industry transformation. This note analyzed the fast fashion industry’s unethical labor routines, environmental implications, and intellectual property obstacles. Efforts have been put in place, but increasing enforcement and applying stricter legislation is needed to truly combat the monumental effects wrapped up in the fast fashion industry.

190. Huckle, supra note 179.
The Impact of the Exhaustion of Rights Doctrine on Parallel Imports and International Trade

CHRISTIAN PEDERSON

I. INTRODUCTION

Copyrights, patents, and trademarks are the various intellectual property (IP) rights that brand owners look to protect to safeguard their brand. Third-party resellers look to exploit price discrimination strategies implemented by companies all over the world to make vast profits. The resellers buy goods for a low price in one country and then ship them to a country where the same good costs much more to make a profit. Parallel importers are only able to exploit these opportunities if the country’s exhaustion of rights doctrine allows for it. There are four different approaches to exhaustion of rights followed throughout the world: (1) national exhaustion, (2) international exhaustion, (3) regional exhaustion, and (4) the materially different approach.

International exhaustion is the most popular approach taken by countries throughout the world. Although this approach can provide cost savings for consumers, it disincentivizes companies from creating new IP due to lost profits from parallel imports. This approach loosens the rights of IP owners to encourage unimpaired free trade around the world. Alternatively, by encouraging countries around the world to strengthen IP rights by adopting the national exhaustion approach, companies both small and large will have the incentive to continue creating new IP without the risk of losing substantial profits to unauthorized resellers of their goods.

This paper looks at the different approaches nations take regarding the exhaustion of rights doctrine as it relates to parallel imports. This paper then goes into a brief industry analysis of the personal luxury goods industry that depicts how the rise of e-commerce is helping the gray market grow very quickly due to the internet’s global reach. The paper offers two separate solutions to handle the exhaustion of rights considering two separate scenarios: (1) in the likely event of no global uniform adoption of an approach, to have major trade agencies
recommend a Pigouvian tax on all parallel imported goods and (2) if it was a perfect world and uniform adoption is possible, argue that a national exhaustion of rights regime is the best approach for worldwide adoption. After discussing the solutions, this paper then goes into a special exception for developing nations. This exception would be put in place to ensure that the development of their economies would not be hindered regardless of what approach is taken.

II. TREATMENT OF PARALLEL IMPORTS

Today, there has been an ever-increasing trend towards globalization and the international free flow of goods. This trend, while good for consumers, creates far-reaching problems for companies through the unauthorized importation of products into countries, which is commonly known as parallel importation. "Parallel imports, also known as gray-market goods, are products made or authorized by the trademark owner for sale abroad, which are then sold . . . back into the domestic market without the consent of the trademark owner" by an unauthorized entity. Parallel importing is the process by which products that are legitimately manufactured and made for sale in Country A, under license from the IP owner at an agreed price, and are then subsequently exported for resale at a set lower price into Country B without the authorization of the licensor or their exclusive licensee.

There are four exhaustion of rights approaches that countries take when it comes to regulating parallel imports, these are: (1) national exhaustion; (2) international exhaustion; (3) regional exhaustion; and (4) the materially different approach. However, there is no international consensus on how to handle parallel imports; both the Paris Convention and the Trade-Related Aspects of Intellectual Property Rights (TRIPS) Agreement do not address it. In terms of the TRIPS Agreement, it simply says that none of its provisions, except those dealing with non-discrimination, can be used to address the issue of exhaustion of IP rights in a World Trade Organization (WTO) dispute. Furthermore, the Doha Declaration clarifies that this allows

---

3. SQUYRES & NORTON, supra note 1, § 29:4.
4. Id.
WTO members to choose how to deal with the exhaustion of rights in a way that best fits their domestic policy objectives. This means that there is no uniform approach that has been recommended by WTO for its member countries.

The most common reason for parallel importation is the large number of profits that can be made by exploiting price discrimination strategies employed by companies throughout the world. Through this, the parallel importer has an incentive to buy the product at a low cost in a foreign market, and then ship it back to sell it as a third party reseller in the domestic market for a profit. Parallel importers will sell these goods at the highest possible price they can, which leads to only a marginal decrease in price for consumers (ranging between zero and two percent).

Parallel imports pose a unique challenge for IP owners due to disputes as to whether the IP rights owner has exhausted their right to control where the good goes after its first sale. Although differences between the domestic and foreign versions of the parallel imported good may be slight, a consumer buying the parallel imported good may be disappointed by the differences and not buy the product again, resulting in lost future profits for the company that owns the IP rights to the good. With the rise of parallel importation, IP owners and their authorized distributors “began to face not only inter-brand competition from other manufacturers selling similar goods under their own trademarks, but intra-brand competition from unauthorized gray market distributors who sold similar goods under an identical trademark.”

Generally, the markets most prone to parallel imports are markets that: (1) are highly competitive; (2) have a heavy reliance on informational and promotional advertising; and (3) involve sophisticated product lines that devote substantial attention to product quality, warranty protection, and the training of distributors. Parallel imports are widespread across many industries and can include automobiles,

6. Id.
7. See id.
9. Id.
10. Id.
11. Id.
13. Id. at 498–99.
books, video games, pharmaceuticals, luxury goods, and apparel. Some industries have much stronger interests in regulating parallel imports than others. For example, the pharmaceutical industry has a very strong interest in limiting or banning parallel imports, and the regulation of parallel imports in the pharmaceuticals industry has slowly become a critical issue the global trade system must deal with. This is a big issue for the pharmaceutical industry because the parallel importation of pharmaceuticals can have severe consequences, including death, when a drug that is subject to a foreign country’s drug regulations is brought into a market it is not meant for. There is also the risk that if the drug is a parallel import from an unauthorized reseller that it could be a fake or low-quality drug coming from China, India, or another country known for ineffective and harmful products. The main risks associated with purchasing parallel imports generally are the lack of customer support, difficulty enforcing consumer rights, and difference in overall product quality. For example, “[t]he British version of ‘Sunlight’ dishwashing soap produces less suds, and the American version is formulated to work best in the ‘soft water’ available in most American cities, whereas the British version is designed for ‘hard water’ common in Britain.”

Although there are many risks associated with parallel imports, consumers also receive tangible benefits from purchasing them. The two most common benefits are increased product accessibility and price savings. Parallel importation is thus able to generate a consumer surplus and public welfare benefit by providing consumers with lower prices for the same products and making products more accessible to low-income individuals who could not afford them. Where consumers benefit from the lower prices parallel imports produce, the unauthorized importers also benefit from expensive manufacturing and advertising.

17. Id.
campaigns without incurring any of the accompanying costs that the IP owner must bear.\textsuperscript{22}

\textbf{III. Exhaustion of Rights: The Four Approaches}

There are four different approaches to the exhaustion of rights doctrine, these approaches are: (1) national exhaustion; (2) international exhaustion; (3) regional exhaustion; and (4) the materially different approach.\textsuperscript{23} Exchanging products in the global marketplace requires application of the exhaustion of rights doctrine to enhance consumer certainty and predictability surrounding the movement of products within the market.\textsuperscript{24} There is a need for a uniform exhaustion of rights doctrine for IP owners to efficiently employ strategies like price discrimination, as well as to empower innovation with the economic incentive of strong IP rights.\textsuperscript{25} The two considerations influencing the development of a country’s exhaustion doctrine are the relationship between international trade and IP rights, and the role of trade and IP rights in guiding economic development within a nation state.\textsuperscript{26} This section will discuss each of the approaches, look at the benefits and risks of each of the approaches to both consumers and producers, and discuss how approaches differ in nations around the world.

\textbf{A. National Exhaustion}

In countries that have adopted the national exhaustion of rights approach, IP rights are exhausted after the protected goods are put into the market. However, the IP owner can still bar the distribution of parallel imports entering from outside of the country.\textsuperscript{27} From the perspective of an IP rights owner, national exhaustion is the preferred method because it advocates for the strongest IP rights of the different exhaustion of rights approaches.\textsuperscript{28} Under the national exhaustion approach, parallel imports are banned from the country, meaning parallel imports cannot confuse or mislead consumers due to differences

\textsuperscript{23} Squyres & Norton, supra note 1, § 29:4.
\textsuperscript{25} Id.
\textsuperscript{26} Id. at 207–08.
\textsuperscript{27} Id.
\textsuperscript{28} Id.
with domestic products and therefore prevents consumers’ perception of the brand from being harmed by a substandard product experience. The following countries have adopted the national exhaustion of rights approach: Albania, Angola, Bangladesh, Bhutan, Cape Verde, Macedonia, Mauritius, and Nepal.

National exhaustion is an example of a weak exhaustion doctrine, meaning it creates strong IP rights for IP owners to be able to exclude a purchaser’s sale or distribution of a product after a sale occurs. This type of exhaustion policy creates strong economic incentives for innovation, which is a major consideration policymakers seeking economic development in a nation must account for. The issue with adopting the national exhaustion approach on an international scale is that it would be fundamentally inconsistent with the principles of international harmonization of free trade policy, which is meant to break down this type of national division for allocating substantive rights. The world has been trending towards greater free trade for decades, and adopting this standard would be a complete reversal on how the world has trended towards the free flow of goods.

The national exhaustion approach would allow countries to employ price discrimination strategies more effectively around the world because it would allow them to sell products at a lower price in less developed countries without the risk of those products then becoming parallel imports. This is a result of price discrimination encouraging more consumers to be involved in the market because consumers willing and able to pay higher prices will do so, like consumers in the US, and consumers unable to pay the higher prices but are still willing to do so, such as consumers in developing nations, are also able to have access to the good. By preventing products from being able to be bought domestically and then resold in other markets effectively under the national exhaustion approach, IP owners are able to segment the market for those goods geographically. This market segmentation is what allows the IP owner to charge different prices for the same product on a market-to-market basis. Market segmentation also allows IP

29. Id.
30. See id.
31. See Brown, supra note 24, at 208.
32. Id.
34. See Brown, supra note 24, at 208–09
35. Id.
37. Id.
owners to alter the product itself in order to either better meet the needs of consumers or to maximize profits on a product-by-product basis.\footnote{38}

By being able to employ price discrimination strategies under the national exhaustion approach, IP owners are encouraged to continue innovating and creating new products because the certainty of IP protections makes it so they would be able to optimize their return on investment by developing a product.\footnote{39} The tricky aspect of the exhaustion of rights doctrine is that it is difficult to find a middle ground that must satisfy IP owners’ needs for economic incentives to create new products, but not at the expense of the well-being of the market and consumers.\footnote{40} When looking at the national exhaustion approach, it offers the ability to provide strong economic incentives to IP owners while also balancing the need for product accessibility in the global marketplace.

The International Trademark Association (INTA) advocates for a national exhaustion regime to be uniformly adopted worldwide as it recognizes that this approach provides both the strong legal rights and economic incentives for continued innovation and creativity throughout the world.\footnote{41} INTA argues that the national exhaustion approach most appropriately considers many of the brand protection concerns that the international exhaustion approach does not address.\footnote{42} Due to price discrimination by IP rights owners, the prices at which products are sold can vary greatly from country to country for a variety of reasons, including, “differences in regulatory requirements, environmental standards, labor and material costs, and government subsidies and taxes.”\footnote{43} By exploiting these conditions, parallel importers are able to buy products in relatively inexpensive markets and then sell them in another market where prices are higher, usually due to local product differences, contractual necessities, local labor costs, or regulatory requirements.\footnote{44} These conditions are why advocates for national exhaustion believe this is the best approach equipped to protect the rights of IP owners.

\footnote{38. Id.}
\footnote{39. See Brown, supra note 24, at 209.}
\footnote{40. Id.}
\footnote{41. Parallel Imports / Gray Market, INTL TRADEMARK ASS’N, https://www.inta.org/topics/parallel-imports/ (last visited Mar. 18, 2023).}
\footnote{43. Id.}
\footnote{44. Id.}
B. International Exhaustion

The international exhaustion of rights approach operates under the rationale that after the IP owner, or an entity under its control, puts the protected product in commerce anywhere in the world, the IP owner is unable to prevent the resale of that product anywhere in the world.\(^\text{45}\) The rationale for this theory is that: (a) IP rights should not be used to interfere with markets by controlling a product and its future destinations after the product has already been placed into the market; (b) the mark is an indication of origin or source of the goods, not a badge of control over future use and sale of the product; and (c) parallel imports do not affect the function of the IP owners mark as a guarantee of the source and quality of the goods; therefore, they do not harm the IP owner and they do not mislead consumers.\(^\text{46}\) If the international exhaustion of rights approach is strictly applied, an IP owner cannot bar a parallel import from entering the domestic market.\(^\text{47}\) Due to allowing parallel imports from anywhere in the world under this approach, it effectively undermines companies efforts to use price discrimination as a strategy to maximize their profits, which can lead to benefits for consumers.\(^\text{48}\) Some countries that use the international exhaustion of rights approach may bar importation if the domestic product and parallel import are materially different—this approach will be discussed later.\(^\text{49}\) Some of the countries that follow the international exhaustion approach include Argentina, Australia, Brazil, China, India, Japan, Mexico, Thailand, and many others.\(^\text{50}\)

Although these countries allow parallel imports, the IP owner still may have other means to challenge the sale of the parallel import after it enters the country, even if customs will not prevent the entry of the parallel import.\(^\text{51}\) Various options that can be used for challenging parallel imports include laws relating to: contract, consumer protection, unfair competition, labeling, weights and measurement, health, and food and drug.\(^\text{52}\) In countries such as Thailand, there is even the possibility to have action taken against a parallel importer if the IP owner can prove that taxes and import duties are not fully paid on the

\(^{45}\) Squyres & Norton, supra note 1, § 29:6.
\(^{46}\) Id.
\(^{47}\) Id.
\(^{48}\) See Brown, supra note 24, at 208–09.
\(^{49}\) Id.
\(^{50}\) Id.
\(^{51}\) Id.
\(^{52}\) Id.
parallel imported product in accordance with Thai tax laws. There are some countries that do employ an international exhaustion regime but ban the parallel importation of certain products, such as pharmaceuticals. This shows that there are many different avenues that can be taken, even in an international exhaustion of rights approach by IP owners, when it comes to protecting their products and brand against potential damages from the sale of parallel imports.

Although there are many ways that IP owners look to be able to prevent the sale of parallel imports because it harms their bottom line, consumers, especially in developing nations, benefit greatly from the international exhaustion approach. International exhaustion helps to deter against anti-competitive behavior due to the policy of global free trade that it encourages. Many smaller developing countries heavily favor the international exhaustion approach because it provides them greater accessibility to products they would typically have little or no access to. The disparity in market conditions on a nation-to-nation basis is so great that it produces the demand for parallel imports. By employing the international exhaustion approach, a country applies a strong exhaustion doctrine, which empowers consumers and limits the ability of IP owners to control the post-sale conduct of consumers. This policy facilitates the alienability of protected products and free trade in the global and domestic markets. These incentives must be carefully weighed against the risk of deterring innovation due to an IP owner not having reasonable expectations of being able to profit.

C. Regional Exhaustion

In countries that have adopted the regional exhaustion of rights approach, they are members of a regional cooperative, typically based on a treaty, where the member countries agree to allow unimpeded

---

54. Id. at 6.
56. Id.
57. See generally Shi, supra note 2, at 241(supporting the idea that parallel importation allows more people to access products because it lowers their price).
58. See id.
60. Id.
61. Id.
circulation of goods within the specified region. Under the regional exhaustion approach, IP rights are exhausted after the IP owner, or a third-party entity within its control, puts the protected product into the region. Once the goods are put into commerce anywhere in the region, the IP owner has no recourse to prevent further distribution of the product in the region. Although, the IP owner can prevent protected goods from outside of the designated region from entering into the region because they would be considered to be parallel imports. The two regions that apply the regional exhaustion of rights approach are the European Union (EU) and the Caribbean Community (CARICOM).

Although the EU and CARICOM both follow the regional exhaustion approach, the way they operate is very different. The CARICOM treaty was created with the goal of integrating economies, coordinating foreign policies, and cooperating in areas including industrial relations and social security. The CARICOM treaty does not expressly address a regional exhaustion of rights regime. However, it has the effect of creating one because the CARICOM treaty prohibits the imposition of new restrictions on imports and exports of products of community origin by member states and requires member states to eliminate existing restrictions on imports and exports of goods of community origin, and requires member states to grant freedom of transit within the community with respect to goods, vessels, and other vehicles transporting goods.

Before the EU was established, the Treaty of Rome established the free flow of goods through the European Common Market. Today, in the modern EU, nothing can restrict the free flow of goods throughout the region, including IP rights. The European Court of Justice has made many decisions which implement regional exhaustion as the prevailing approach taken by the courts. The European Court of Justice has even gone so far as to rule that “if the goods are in transit through the [European Economic Area] EEA, but not intended for the EEA, there may still be infringement in the EEA if the goods, without authorization, bear a trademark that is identical or nearly identical to a
trademark registered in the EU . . . .”72 Furthermore, Article 7 of Directive 2008/95/EC dated October 22, 2008, mandates that IP owners cannot use their IP rights to prevent the circulation of goods within the community unless the goods are altered, damaged, or enter without the IP owner’s consent.73

Regional exhaustion of rights is an approach that does not provide full protection to IP owners. Since the EU employs a regional exhaustion approach, their approach would not stop an unauthorized distributor from purchasing a product in a country where the good is cheap and then selling it within the region of a country where the good is relatively more expensive. Figure One shows the price level indices (PLIs) for total household consumption of goods and services in the EU in 2020.

Figure One: Price Level Index for Final Household Consumption Expenditure in the European Economic Area in 2020

This figure shows 2020 price levels varied greatly throughout the EU and the European Economic Area (EEA) nations that also participated in the study. In this study, the EU average PLI is set at 100, which means that over half of the countries in the EU (primarily the Eastern EU) have prices that are below the average price. This figure shows that on the low end of the spectrum, prices in Turkey are on average fifty-six percent lower than the EU average, whereas prices in Switzerland are seventy percent above the EU average.75 Pricing disparities this great

73. Squyres & Norton, supra note 1, § 29:8.
75. Id.
within Europe allow significant opportunities for resellers within the region to exploit the pricing differences between low and high price countries. This sort of scenario worries IP owners and spurns them to advocate for a national exhaustion approach because it is very possible for them to lose out on profits in countries where the prices are higher due to unauthorized distributors selling the IP owner’s same product for a lower price.

D. Materially Different

Some countries that employ either the regional or international exhaustion approaches have exceptions to allow parallel imports to come into the domestic market if the parallel imported good is materially different from the good sold in the domestic market. A material difference is anything that may mislead or confuse a consumer, including any difference that has the ability to impact a consumer’s willingness to purchase the product or leads to dissatisfaction with the product after it is purchased. The rationale behind this exception is based on protecting the consumer from being misled or confused about the product the consumer is purchasing and protecting the reputation of the IP owner. Jurisdictions that use this approach include the EU, Canada, South Africa, Sri Lanka, Switzerland, and the United States.

What is considered to be a material difference varies widely country-by-country, but there are some material differences that are common amongst regions that follow this approach. This includes the parallel import not complying with domestic regulations; the packing or instructions not being in the primary language of the domestic market; the parallel import tasting, smelling, or looking different because it was made for foreign market preferences; or if domestic conditions could impact product performance. Each of these constitute a material difference from the domestic product because of how they stand to substantially alter the consumer’s experience that they are expecting to have.

One of the main reasons for the existence of the material differences standard is that in some countries there are political or other considerations that make it impracticable for countries to shift

76. Squyres & Norton, supra note 1, § 29:10.
77. Id.
78. Id.
79. Id.
80. Id.
from international to national exhaustion.\textsuperscript{81} It is in these sorts of countries that the material differences approach is needed because it provides for the exclusion of parallel imports that are materially different from the protected goods authorized for sale in the domestic market by the IP owner. It must also be noted that material differences need not be physical in nature. Many non-physical differences may be material to the consumer, such as the absence of product support, warranty, or instructions in the consumer’s native language.\textsuperscript{82} Similarly, the differences in question do not need to threaten the health or safety of a consumer because the consumer’s legitimate expectations of the product may be let down regardless of the risk of any physical harm.\textsuperscript{83} Though it may not be considered a relaxed standard by any means, countries adopting this approach afford IP owners the opportunity to differentiate their products for different markets so they will be protected against parallel importation.

\textit{E. Enforcement Mechanisms}

Many countries that apply a national or regional exhaustion approach allow for the recording of IP registrations with customs officials so that they will block the infringing products from entering the country or region.\textsuperscript{84} Recording an IP registration with customs can be a very effective way to stop parallel imports at the border rather than waiting to go through the country’s court system to attempt to stop an infringing product that has already been introduced to the market.\textsuperscript{85} In Cambodia, an IP owner has the ability to register exclusive distributorships they have within the country with the Department of Intellectual Property Rights (DIPR).\textsuperscript{86} Once the exclusive distributorship has been registered, it is published for three months to inform the market, and once this period has ended the IP owner or distributor may start enforcing against parallel imports.\textsuperscript{87} This sort of enforcement mechanism is effective at enforcing rights against parallel imports because it puts the market on notice that enforcement on those goods will be occurring, starting at a specified date. Without this

\textsuperscript{81}A “Material Differences” Standard for International Exhaustion of Trademark Rights, supra note 42.
\textsuperscript{82}Id.
\textsuperscript{83}Id.
\textsuperscript{84}SQUIRES & NORTON, supra note 1, § 29:12.
\textsuperscript{85}Id.
\textsuperscript{87}Id. at 5.
registration, an IP owner is not able to enforce their rights against parallel imports because the market has not been put on notice.

If an unauthorized reseller alters a good, an IP owner can look to enforce their rights, as it may be considered trademark infringement. If an unauthorized reseller alters a good, an IP owner can look to enforce their rights, as it may be considered trademark infringement. Trademark infringement mainly encompasses protecting the trademark owner’s reputation and goodwill against being damaged in the eyes of consumers due to unauthorized use of the trademark. An example of alteration of a good that may cause consumer confusion, and thus equate to trademark infringement, is the Lil Nas X “Satan Shoes” lawsuit that was brought by Nike. This lawsuit revolved around customized pairs of the Nike Air Max 97 sneakers that were said to contain a drop of human blood in the midsoles, and are printed with Luke 10:18, a bible passage that refers to Satan’s fall from heaven. In this lawsuit, Nike was able to successfully convince the judge that “even ‘sneakerheads’ were actually confused . . . ” by the shoes because of the prominent Nike “swoosh” logo. Nike was able to show that the altered good had a high likelihood of confusing consumers and damaging their brand reputation internationally. However, it is up to brands to determine when they want to enforce their rights for trademark infringement. In 2019, the same company that made the “Satan Shoes” made “Jesus Shoes” based on the same Nike shoe model, and Nike made no objections.

As has been shown, there are various options that IP owners have to enforce their rights against parallel imports even in a country that takes an international exhaustion approach. Some countries implement fair trade and unfair competition laws that may be used to exclude parallel imports from the market. Unfair trade arguments encompass situations where parallel imports may: (1) interfere with contracts with domestic distributors; (2) mislead the public; (3) take unfair advantage of the advertising and marketing efforts of IP owners; and (4) injure or restrain a domestic industry or business. In countries following UK law, the only unfair competition claim an IP owner can make is a

88. Id. at 11
89. SQUIRES & NORTON, supra note 1, § 29:13.
91. Id.
92. Id.
93. See generally id. (this is shown through the Court’s granting of the restraining order).
94. Id.
95. SQUIRES & NORTON, supra note 1, § 29:14.
96. Id.
passing off claim. In order to establish a passing off claim, the IP owner must show: (a) goodwill associated with its mark; (b) a misrepresentation as to source, usually due to another trader using the mark in the course of trade; and (c) a likelihood of damage to goodwill as a result of the misrepresentation. In order for a passing off claim to prevail, the IP owner “will probably have to establish that there is a difference in the quality of the goods that will mislead the potential consumers and damage the trademark owner’s goodwill associated with the mark.” The IP owner may also be able to argue a passing off claim if the parallel importer passes themself off as an authorized dealer.

IV. IMPACT ON GLOBAL TRADE: E-COMMERCE

Parallel imports have a large impact on the global economy, and they make up a significant amount of yearly total sales within various industries. This has been amplified by the growth of e-commerce, and how the internet has given parallel importers the unprecedented ability to sell protected goods in foreign markets by merely advertising lower prices on protected goods coming from outside of the market to entice consumers to order parallel imports. In a survey conducted by KPMG, it was estimated that in 2009, the gray market cost IP owners up to $63 billion in US sales alone. Further, in 2011 it was estimated that tech companies such as Samsung and Hewlett-Packard lose up to $1.4 billion a year in global sales.

One of the industries that is most prevalent to goods becoming parallel imports is the personal luxury goods market. In the personal luxury goods market, it was estimated that the gray market is worth up to eight percent (roughly $20.56 billion) of the $257 billion global market. It is very common for purchasers who are buying gray market goods (parallel imports) to receive a discount, usually ranging

97. Id. § 29:15.
98. Id.
99. Id.
100. Id.
102. Id.
103. Id.
105. Id.
from fifteen to thirty-five percent off of what buying the good from an authorized distributor in the region would cost. Consumers turn to parallel imports for these sorts of goods because the price that they are regularly available for in the domestic market is too high for a large number of consumers. E-commerce companies such as Baltini and Italist have quickly become major companies for selling the latest fashion styles at a steep discount. Cettire, a gray market fashion company, was listed on the Australian Stock Exchange at the end of 2020. Cettire has a stock price of $3.72 per share with a market cap of $1.42 billion as of December 20, 2021. These are only a few of the vast number of companies that have large-scale gray market distribution operations that have emerged as players in the market as of recently.

E-commerce companies like Cettire exploit legal loopholes to use price discrimination to their advantage. Due to luxury fashion brands wanting to have an exclusive feel to their products (i.e., not many people can afford them so there is a lower supply), parallel importers are able to offer the product to consumers who have less spending power but still have the willingness to purchase the good. On Cettire, there are a large number of products that are on sale for substantial discounts from the suggested retail price (e.g., Bottega Veneta slides were twenty-four percent off, and a chain pouch leather bag was on sale for $300 off its recommended price). Cettire is able to do this because “particularly in Southern Europe, genuine product has often had a way of coming out the back door of a factory or off the back of a truck and eventually into the hands of consumers looking to pay less than full price.” Companies like Cettire are able to get genuine products for lower prices from other nations and illicit sources, and then by exploiting import duty laws (e.g., shipments valued less than $800 are generally shipped free of import duties into the US), they are able to bypass certain regulations that make the cost of products higher domestically.
V. THE SOLUTION: HOW SHOULD PARALLEL IMPORTS BE DEALT WITH?

The issue presented to nations and regions is what type of exhaustion doctrine should they apply to most effectively meet the needs of their nation when balancing both IP owner and consumer interests. This section presents solutions for two separate scenarios. Firstly, in the likely event that uniform adoption is infeasible, countries should urge the WTO and other governing world trade agencies to encourage countries that currently use an international exhaustion regime to adopt a Pigouvian tax on all parallel imports. Secondly, if uniform adoption is feasible, countries should adopt a national exhaustion approach to incentivize domestic innovation and safeguard consumers from unfair trade practices. However, it would be extremely difficult to implement a national exhaustion regime in many developing countries, so exceptions regarding developing countries with a legitimate need for a weak exhaustion doctrine would be necessary.

When looking at parallel imports from an economic lens, it is evident that a nation’s economy needs to reach a certain stage of overall development before it has the ability to commit a substantial amount of its resources to research and development and initiate genuine efforts to protect IP rights.115 In many developing countries, in order “to harness the weak purchasing power of low-income citizens . . . authorities of developing countries often turn a blind eye on counterfeiting and piracy because it is a ‘legitimate grey market’ in the early stages of development as they know, they need time to phase-in effective policies.”116 This sort of policy in developing countries also allows the governments to maintain the national revenue that is generated from pirated and imitated goods.

A. No Uniform Adoption: Pigouvian Tax

In the scenario where global uniform adoption of an exhaustion system is not possible, countries implementing an international exhaustion or material differences approach should be encouraged to adopt a Pigouvian tax on parallel imports. A Pigouvian tax is assessed against private individuals or businesses that engage in activities that create adverse effects for society.117 Adverse side effects are costs that are not included in the product’s market price (e.g., it can include

115. Shi, supra note 2, at 231.
116. Id. at 230.
pollution, strain on healthcare from tobacco sales, parallel imports deterring innovation, and any other side effects that have an external, negative impact on society as a whole). The rationale behind this sort of tax is that it redistributes the cost back to the producer or user of the negative externality in order to deter the conduct. Although Pigouvian taxes are meant to equal the exact cost of the negative externality in theory, this can be very difficult to determine and if overestimated, there is a risk of harming society instead of creating any tangible benefits.

To understand this solution, externalities must also be discussed before going further into the proposed solution. An externality is either a cost or benefit caused by a producer that is not financially incurred by that producer. The costs and benefits arising from externalities can be both private (i.e., to an individual or private organization) or social (i.e., to society as a whole). Externalities typically occur when the production or consumption of a certain good or service has an impact on a third party that has no direct relation to production or consumption of that good or service. In the situation of parallel imports, they are considered to be technical externalities, which means the indirect effects of the externality have an impact on the consumption and production opportunities of others, but the price of the product does not take those externalities into account. Parallel imports create a negative externality due to the costs it imposes on IP owners not only financially but also socially due to how the lack of protection can lead to less innovation. Thus, in order to overcome externalities, there are three solutions that can be taken, which are: (1) taxes (e.g., to reduce the negative effects of pollution); (2) subsidies (e.g., subsidize orchards that plant fruit trees to provide positive externalities to beekeepers); or (3) governmental regulations (e.g., environmental regulations and health related legislation).

When looking at the specific problem caused by parallel imports, it is important to consider the benefits and risks to parties on all sides of the spectrum. From an IP owner’s perspective, they have a substantial

---

118. Id.
119. Id.
120. Id.
122. Id.
123. Id.
124. Id.
125. See generally id. (supporting the idea that government regulation of parallel imports can help offset the effects of externalities).
126. Id.
interest in strong protection of their IP rights due to the expectation of being able to maximize profits from their IP creations.\textsuperscript{127} This is because in sectors such as high-technology markets (e.g., pharmaceuticals), there is an extremely high risk of failure of expensive research on any given product they attempt to develop.\textsuperscript{128} Therefore, companies that operate in these high-tech sectors must have a reasonable expectation that the profits of endeavors that do succeed will be maximized in order to cover the cost of failed projects.\textsuperscript{129} If nations fail to provide protections to IP owners, innovative activities by high-tech companies are at risk of drying up.\textsuperscript{130} It has been noted by prominent economists Nancy Gallini and Suzanne Scotchmer that there is the risk of the balance between private gain and public good slipping too far one way or the other beyond what is necessary to encourage innovation.\textsuperscript{131} This in turn can become a drag on innovation that leads to the impediment of creation of further innovations by individuals and companies.\textsuperscript{132} The World Health Organization and the World Bank have gone as far as to suggest that countries and world organizations should give rewards for developing vaccines that would not be developed or might not be widely distributed under a proprietary rights system.\textsuperscript{133}

Although there are substantial benefits that come with strong IP rights for IP owners, there are also risks that come with employing too strong of IP protections in a country.\textsuperscript{134} From a developing nation’s standpoint, employing strong IP protections may lead to companies becoming monopolies, leading to higher prices.\textsuperscript{135} Strong IP protections may also hinder rather than stimulate trade, which can act to the detriment of global development in the long run because trade is the primary way in which technology transfer is achieved.\textsuperscript{136} This creates the need for developing nations to carefully construct an IP system that balances IP protection and exploitation.\textsuperscript{137} Thus, when constructing a Pigouvian tax, it is important to take into consideration the impact such a tax would have on developing nations in order to avoid hindering developing economies.

\begin{itemize}
\item[\textsuperscript{127}] See Shi, supra note 2, at 223–24.
\item[\textsuperscript{128}] Id. at 224.
\item[\textsuperscript{129}] Id. at 223–24.
\item[\textsuperscript{130}] Id. at 224.
\item[\textsuperscript{131}] See Nancy Gallini & Suzanne Scotchmer, Intellectual Property: When is it the Best Incentive System?, 2 INNOVATION POLY & ECON. 51, 51 (2002).
\item[\textsuperscript{132}] Id.
\item[\textsuperscript{133}] Id. at 55–56.
\item[\textsuperscript{134}] Shi, supra note 2, at 230.
\item[\textsuperscript{135}] Id. at 231.
\item[\textsuperscript{136}] Id. at 230–31.
\item[\textsuperscript{137}] Id. at 230.
\end{itemize}
From the consumer’s perspective, there are also benefits and risks that must be weighed when it comes to determining whether a Pigouvian tax is the most efficient way to deal with the negative externalities created by parallel importation. Consumers have a strong interest in weak IP rights protections because weaker protections allow them the opportunity to purchase products that they would normally not have the ability to afford. This comes from the greater accessibility and lower prices that parallel imports are able to afford consumers in both developed and developing nations. This policy of allowing parallel imports to benefit consumers for the aforementioned reasons is evident in Indonesia’s policy on pharmaceutical parallel imports. Indonesia is unique because it does not recognize a national exhaustion doctrine, and parallel imports are strictly limited to pharmaceutical products under Patent Law No. 13 of 2016. This ensures that pharmaceutical products are reasonably priced and widely accessible to all Indonesian citizens. Indonesia employs this policy to make sure there is more competitive pricing in the market and to help minimize the risk of a monopoly in the pharmaceutical sector because parallel imports broaden the public’s access to pharmaceutical products.

Though parallel imports do improve the purchasing power of consumers and allow greater access to items, parallel imports also can be detrimental because of the parallel import being altered, damaged, or materially different in a way that harms the consumer’s perception of the brand. A consumer’s bad customer experience, even if it was with a parallel import, has a direct correlation with a reduction in a company’s sales. Thus, as stated previously in this section, it is important to protect the expected profits of an IP owner in order to encourage them to continue to innovate without the fear of significant lost profits in the future. Not doing so can actually end up hurting the consumer as a result of company’s either increasing prices or cutting production as ways to deal with the parallel import problem.

139. Id. at 208.
140. Adcock, supra note 86, at 6.
141. Id.
142. See id.
143. See generally Shi, supra note 2, at 224 (supporting the contention that if companies cannot expect to make profits to cover their costs in certain areas they will eventually end business activities in certain markets).
When looking at all these factors together, the best solution to ensure that both consumers and IP owners are protected is to implement a Pigouvian tax that is aimed at deterring parallel importation from occurring. This Pigouvian tax is meant to act like an import duty that is meant to favor domestic products. For this Pigouvian tax to be effective, it should not be framed in a way that punishes the producers to a greater extent than the harm they are causing to IP owners. This will require the Pigouvian tax to be applied on a country-by-country basis with a set of frameworks about how to calculate the amount that should be charged to counteract the negative externality.

Due to the need for a government to know a near exact amount of the externality cost produced by parallel importers, a worldwide database of suggested retail prices for protected goods in each participating nation would need to be established. This will also require each country to establish a customs registry database of all protected goods so that officials will be able to quickly and efficiently find out if a good is protected or not. Once globally suggested retail prices and global registry databases have been established, there will be a worldwide ability to calculate a Pigouvian tax meant to counteract the profits that parallel importers intend to make. This tax is primarily meant for large wholesale-type shipments to deter large-scale parallel importation.

If an unauthorized distributor attempts to bring bulk shipments of parallel imports into a country, the Pigouvian tax will be implemented by having the database find the suggested retail price in the exporting country and comparing it to the suggested price in the importing country. The percentage difference between these prices will then be charged as an import tax or duty if the price in the importing country is higher than the exporting country. This tax will deter parallel imports because parallel importers will lose money in their endeavor. They will end up losing money because the implementation of a tax that covers the price difference between countries as well as the added costs of international shipping will make parallel importers have to charge more for the product than the IP owner does. In this scenario, even in the event an individual or business decides to engage in parallel importation, they will most likely not have any effect on an IP owner's expected profits. By implementing a Pigouvian tax that acts as an import tax or duty, nations will be able to help fund their IP protection programs, as well as other business sectors which will in turn help their economies grow. However, there is a special exception for developing nations that is discussed below.
B. Uniform Adoption: National Exhaustion

In the event that there were perfect conditions around the world to implement one exhaustion system, a national exhaustion regime would be the most appropriate solution when taking into account the needs of IP owners and consumers. National exhaustion affords IP owners strong rights while simultaneously protecting consumers from goods that do not come from an authorized source that may be damaged, altered, or materially different in a way that would harm their perspective of the product/brand.\textsuperscript{148} National exhaustion is the best approach to continue to encourage innovation and new technologies to be developed. To do this, countries should require IP owners to register their products and exclusive distributorships with customs in each country. In this way, customs officials would be able to accurately tell if a good is coming from an authorized distributor or an illicit third party.

Without the ability to reap the fruits of their labor, inventors have little reason to innovate, so it makes sense to adopt the approach that will keep pushing the world towards new types of innovative, highly sought-after, products. By allowing companies to employ the business strategy of price discrimination without the worry of having to compete against themselves, they are less likely to raise prices or lower their production because they will not have to compete against parallel imports. This also will encourage companies in high-tech sectors, such as pharmaceuticals, to continue embarking on ambitious projects that cost significant amounts, even though they have a high risk of failure.\textsuperscript{149} By applying a national exhaustion regime with a special exception that is discussed in the next subsection, there would be an efficient balance between the consumers’ and IP owners’ interests that would foster new innovation and the development of economies.

Though critics may argue that a regional or international exhaustion regime would be more effective, these types of regimes would hinder worldwide IP development. If an international exhaustion regime was uniformly adopted around the world, the rate of parallel importation would skyrocket, and many companies would have to make difficult decisions when thinking about the benefits of creating new products. This is due to the possibility of large amounts of lost revenues because of the abundance of parallel importation that would be expected to go on. Price discrimination strategies that businesses use to price products differently in separate nations would become ineffective. This in turn would harm consumers because they would begin to see lower

\textsuperscript{148} See Squyres & Norton, supra note 1, § 29:4–5.
\textsuperscript{149} See Shi, supra note 2, at 245–47.
product accessibility and higher prices, thus harming economies across the globe. A regional exhaustion regime would also be harmful because it would break the world into however many regional regimes make up the world and would essentially allow parallel imports from only a certain number of countries.\textsuperscript{150} A uniform regional exhaustion approach would also render price discrimination ineffective because purchasing power can vary greatly even in the same region (e.g., the vast disparities in PLIs between European Economic Area countries in Figure 1).\textsuperscript{151} This would also likely lead to decrease in production or increase in global prices in order to combat the parallel importation, which in turn harms the consumers. By adopting a uniform national exhaustion approach, consumers would not be the ones bearing the burden of the cost of lost profits due to parallel imports.

C. Developing Nations Exception

Under both solutions that have been discussed to help solve the issue posed to governments by parallel imports, there is the question of how to handle exhaustion of rights in developing nations. In the global IP system, developed countries are primarily producers of protected goods whereas developing countries are generally consumers of protected goods.\textsuperscript{152} Corporations have the ability to exercise monopoly power that can place “low per capita income countries and smaller economies at a significant disadvantage.”\textsuperscript{153} This requires the international community to take a constructive approach to ensure that measures to enforce IP rights do not obstruct the development of these economies.\textsuperscript{154} In many underdeveloped nations, prohibiting parallel imports could make it nearly impossible to provide consumers in these nations affordable access to quality products.\textsuperscript{155} When looking at pharmaceutical products, it is paramount to facilitate timely access to generic versions of medicines from developed nations via parallel imports in order to introduce competition to brand name drugs. This in turn will force prices to go down to be more competitive in the market, bringing benefit to consumers in these nations.\textsuperscript{156}

In order to refrain from hindering developing nations’ economies while still protecting IP owners from parallel imports, a special

\textsuperscript{150} Squyres & Norton, supra note 1, § 29:4, 7.
\textsuperscript{151} Eurostat, supra note 74.
\textsuperscript{152} Shi, supra note 2, at 248–49.
\textsuperscript{153} Id. at 249.
\textsuperscript{154} Id. at 240.
\textsuperscript{155} Id. at 246.
\textsuperscript{156} Id.
exception must be made to grant developing countries the benefits of favorable treatment that furthers the development of the economy.\textsuperscript{157} This will be done by allowing developing nations that are under a certain GDP threshold to engage in parallel importation until the economy has reached a point that IP rights need stronger protection in those countries. This exception will allow developing nations to benefit from parallel imports up until the point that their respective governments can start making substantial investments in IP and industry.\textsuperscript{158} This exception for parallel imports will be formulated as a “fair use” doctrine for developing economies that encompasses all relevant IP protections. In determining whether a good being imported into a developing economy should be subject to fair use in that nation, there are three relevant factors that should be considered: (1) the purpose and character of the use, including whether the use is of a commercial nature or is for healthcare/nonprofit/educational purposes; (2) the nature of the protected good; and (3) the effect of the use upon the potential market or value of the protected good.\textsuperscript{159} If a good is deemed as being subject to fair use after this three-pronged inquiry, then it should be allowed as a parallel import in developing economies to help them grow. This can be applied to the Pigouvian tax as well because, in scenarios where the protected good passes the “fair use” test, no tax would be applied to the parallel import in developing countries.

Nearly every country passes through a stage of development when the imitation of foreign products is a legitimate and strategic business practice.\textsuperscript{160} Even the United States had a stage in its economic growth largely based on copying technology which culminated in the adoption of technological, economic, and legal policies from England and France.\textsuperscript{161} Thus, allowing for an exception for developing nations only follows the trend that developed nations have been taking for centuries when it comes to the formation of IP protections.

\textbf{VI. Conclusion}

Parallel imports are an ever-growing issue in the world today, and the lack of a uniform exhaustion of rights approach leaves IP owners in

\begin{itemize}
\item \textsuperscript{157} Id.
\item \textsuperscript{158} Id. at 230.
\item \textsuperscript{159} U.S. Copyright Office, More Information on Fair Use, https://www.copyright.gov/fair-use/more-info.html (last visited Mar. 18, 2023) (using factors from Section 107 of the Copyright Act in order to create a “fair use” test that can be applied to parallel imports going into developing nations).
\item \textsuperscript{160} Shi, supra note 2, at 231–32.
\item \textsuperscript{161} Id. at 232.
\end{itemize}
constant doubt as to their IP rights. The rapid and continuing growth of e-commerce has only compounded the problem of parallel importation and made it worse due to how easy it is to ship goods to yourself internationally with the mere click of a button. By adopting either a Pigouvian tax to pass on the costs of the negative externalities created by parallel imports or a uniform national exhaustion of rights approach, countries would be able to effectively and efficiently protect both IP owners and consumers from the harms caused by parallel importation. By implementing these policies as solutions to the parallel importation issue, nations will be able to sow the seeds of innovation, as inventors will be enabled to come up with new ideas without fear of others profiting off them.
Ensuring Dignity for the Survivors and the Dead: Genocide Denial and the Law in Post-Conflict Bosnia and Herzegovina

DANIEL SCHUMICK

"The murdered are to be cheated out of the single remaining thing that our powerlessness can offer them: remembrance." – Theodore Adorno, *Critical Models*¹

“You can imagine that feeling when you walk by this person, and you know what he did, and he knows what he did, and he’s proud. You simply die all over again, all of those feelings from the war come back to you, as you stand next to him in a queue to pay the water bill.” – Munira Subašić, president of the Mothers of Srebrenica²

The Holocaust was undoubtedly one of the most disturbing events in human history. The genocide of about six million Jews revealed a depravity in the human condition that seemed unfathomable. Yet, the Holocaust was not the last genocide to take place on European soil in the twentieth century. The world largely stood by as the Republic of Bosnia and Herzegovina plunged into a deadly war characterized by a brutality not seen since the Holocaust. The slogan “never again” that appeared after the Holocaust evidently fell on deaf ears.

From 1992 to 1995 the Bosnian-Serbs, Bosnian-Croats, and Bosniaks³ fought a war that killed 96,000 people and displaced two

---

³ A quick note on ethnicity, Bosnia is home to three ethnic groups. The southwest is home to a significant Croat population, the west and north have a significant Serb population, and Bosniaks (Muslim Bosnians) live all throughout the country. This paper will use Bosniak to refer to the Muslim population of Bosnia since Bosnian could refer to a Bosnian national of any ethnicity. This paper will use Bosnian-Serb and Bosnian-Croat when the context requires clarity that the group I am speaking about are Bosnian.
million people. The Dayton Agreement stopped the killing and created the new republic of Bosnia and Herzegovina, a unique country comprised of two entities divided along ethnic lines (the Federation of Bosnia and the Republika Srpska). A number of war criminals were tried by the International Court of Justice (ICJ) resulting in a number of war crime, mass rape, and genocide convictions.

Less than thirty years later, the war's social impact is still felt across Bosnia and Herzegovina in the day-to-day lives of those who survived. Ethnic tensions between Bosniaks, Bosnian-Serbs, and Bosnian-Croats continue to plague the country’s political scene, education system, and other day-to-day interactions. Perhaps most notable is the persistence of genocide denial in the Republika Srpska. Despite the ICJ ruling, sixty-six percent of Bosnian-Serbs in the Republika Srpska deny that Bosnian-Serbs committed genocide during the Bosnian war. Denial of the Bosnian genocide and even genocide triumphalism have become mainstays of the Serb member of the Presidency of Bosnia and Herzegovina, Milorad Dodik.

To combat rampant denial of atrocities committed during the Bosnian war and to support social order in the divided country, Valentin Inzko, the High Representative for Bosnia and Herzegovina, promulgated a law that declared public denial of genocide, crimes against humanity, or war crimes illegal. This criminalization of speech would likely catch American observers off-guard. After all, American jurisprudence has made herculean efforts to protect free speech. However, Austria, Belgium, Czechia, France, Germany, Liechtenstein, Lithuania, Luxembourg, Portugal, Poland, Romania, and Switzerland all have laws that criminalize denial of genocide or the Holocaust specifically.

The new genocide denial law has caused an uproar in both the

nationals. I will use Croat and Serb frequently in the historical section because a sizeable portion of the paramilitary groups involved in the Bosnian war were made up of fighters from Serbia and Croatia.

5. See id. at 160.
6. Id. at 161.
8. Code Criminal art.145a (Bosn. & Herz.).
domestic Bosnian community as well as the international community. At first glance, the contentiousness of the law should be obvious. As a general principle, liberal-democratic societies do not criminalize free speech. Rather, liberal-democratic societies tolerate speech of all kinds in the interest of not restricting the essential human right of freedom of speech. However, when perpetrators and victims must not only live side-by-side but also must work together to keep their government functional, the clash between human dignity and freedom of speech becomes more difficult to resolve. The debate is a difficult one and can be a highly academic exercise at times, but these kinds of jurisprudential debates are important because they reflect real social and political conflicts in society.

The Bosnian genocide denial law is also surrounded by procedural controversy in addition to substantive controversy. The man who promulgated the law, Valentin Inzko, is neither a citizen of Bosnia and Herzegovina nor an elected official. Rather, Inzko is an Austrian holding the unelected position of the High Representative for Bosnia and Herzegovina. The High Representative for Bosnia and Herzegovina is appointed by a committee of international actors from the states who oversaw the 1995 peace process in Bosnia and Herzegovina. The problem with an unelected official creating a law that curtails the individual right to speak about a sensitive political/social issue should be obvious. Liberal-democratic societies typically do not legislate without popular consent.

In response to the Bosnian genocide denial law, the Republika Srpska elite, led by Milorad Dodik, have resisted the application of the law and threatened secession. Dodik has demanded greater autonomy for the Republika Srpska, repealing of the genocide denial law, and the
The creation of Bosnian-Serb-only institutions such as a Bosnian-Serb army.\textsuperscript{15} The Republika Srpska’s demands have raised cause for alarm in the West not only for their potential to split Bosnia and Herzegovina but also because of the vocal support Dodik and the Republika Srpska have received from right-wing authoritarian regimes.\textsuperscript{16} The looming crisis in Bosnia and Herzegovina has stirred great concern in the international community and EU officials have even begun to “correct” the genocide denial law to stabilize the situation.\textsuperscript{17} However, doing so would be a mistake.

This note will defend Bosnia’s new genocide denial law by responding to common criticisms of speech regulation laws as well as the manner in which the law came to be. Freedom of speech is not an absolute right and must yield to human dignity when the two conflict. The international community must stand by the law to defend human dignity and the legitimacy of international law from right-wing authoritarian states who seek to revise the liberal world order.

This essay will proceed in five parts: (1) a discussion of the prevalence of outsider scholarship when it comes to academic study of the Balkans; (2) an overview of the historical background that touches on the Bosnian war, the government structure in Bosnia and Herzegovina, and the genocide denial law; (3) criticism of the genocide denial law along with responses to the criticism; (4) the political impact of the law; and (5) concluding remarks.

I. Outsider Scholarship

I would be remiss if I failed to acknowledge the almost suffocating prevalence of outsider scholarship concerning the Balkans. The Balkans, in particular Bosnia and Herzegovina, are often fetishized as an “exotic” part of Europe and thus suffer from orientalist scholarship.


The image of the Balkans as “exotic” often results in western academics and other writers not engaging, or even acknowledging, the work of Balkan academics. Western ignorance, or even flat-out dismissal, of Balkan academic work has caused the emergence of an insider versus outsider debate within Balkan scholarship.\(^{18}\)

As a general matter, the insider-outsider debate focuses on who is in a better position to study a subject, scholars outside the community that is the subject of study or scholars within the community that is the subject of study. Insiders see outsider scholarship as colonial in nature “which ignores, silences, and/or diminishes insider perspectives.”\(^ {19}\) Sociologist Robert Merton describes the insider way of thinking as “only members of a particular group possess the ability to undertake research of their group.”\(^ {20}\) On the other hand, outsiders view themselves as being in a better position to provide an objective account of their research.\(^ {21}\) Sociologist Georg Simmel argues that the outsider is freer, practically and theoretically, to study a given topic because the outsider is less tied to their prior assumptions and knowledge.\(^ {22}\) Then there are some scholars, such as Melissa Gilbert, who argue the debate is unproductive because the very act of research puts “an ‘insider’ in an ‘outsider’ position.”\(^ {23}\)

Gilbert makes a good point about the usefulness of the insider-outsider debate, but the insider critique that outsiders silence or ignore insider voices is also well-taken and easily seen in Western discourse about the Balkans. Ameliorating the outsider problem and recognizing that my voice cannot and should not supplant the voices of those whose experiences I write about, I have taken care to include Bosnian scholarship when available and appropriate.

---

18. A good example of the insider-outsider debate as applied to the Balkans can be seen in the critical reactions to the University of Vienna’s conference on the Srebrenica Genocide. The conference largely excluded Bosnian scholars in favor of German scholars. This drew massive criticism on social media from Bosnian scholars holding a variety of views about the role of outsider scholarship. One group found the debate silly and unproductive. Another group advocated for a larger focus on insider scholarship to allow victims to speak for themselves. Yet another group argued that outsiders can, and should, write about the genocide in Bosnia but should support, not supplant, the voices of those who lived through the experience.


20. Innes, supra note 19, at 442.

21. Id. at 443.

22. See id.

23. Id.
II. HISTORICAL BACKGROUND

Bosnia and Herzegovina was once a republic within the Socialist Federal Republic of Yugoslavia (1945–1992). Yugoslavia was well known as a multiethnic society and was often considered the crossroads of the East and West due to the large presence of Orthodox, Jewish, Catholic, and Muslim religious communities. Bosnia and Herzegovina was considered a microcosm of Yugoslavia due to it having the largest presence of all the major Yugoslav ethnicities (Bosniak, Croat, Macedonian, Montenegrin, Serbian, and Slovene). The Yugoslav state was relatively stable during the presidency of Josip Broz Tito (1945–1980); however, after Tito’s death, the economy collapsed and ethnic tensions rose. In 1991 the republics of Slovenia and Croatia seceded from Yugoslavia, and in 1992 Bosnia and Herzegovina voted through a public referendum to secede as well. The secession of Bosnia, and their large Serbian minority, was not well received by either the Serbs or the Bosnian-Serbs. With the success of the referendum, the Serbs began a war that would tear the newly independent Bosnian state in two.

A. The Bosnian War

In March of 1992, dozens of paramilitary groups flooded into Bosnia and Herzegovina from neighboring Serbia.24 In what would later become known as the Drina campaign, Serb paramilitary groups such as Arkan’s Tigers, the White Eagles, and the Yellow Wasps unleashed a campaign of terror across eastern Bosnia and Herzegovina.25 Their goal was the complete eradication of Bosniaks from what they considered to be rightful Serbian land.26 Vojislav Šešelj, leader of the White Eagles, began his campaign with the following, “[d]ear Četnik brothers, especially you across the Drina River, you are the bravest ones. We are going to clean Bosnia of pagans and show them a road which will take them to the east, where they belong.”27

The paramilitary groups began their genocidal campaign with Zvornik, a city in the northwest corner of Bosnia and Herzegovina near the Serbian border. The Serbs began by using artillery to shell the town throughout the night.28 The next morning, they invaded the town. Serb paramilitary groups killed Bosniak men and boys, raped Bosniak

25. See id.
26. See id. at 112.
27. Id.
28. Id. at 114.
women, looted Bosniak houses, burned Bosniak buildings, and defiled the corpses that they left in the streets.29 Survivors of Zvornik later recounted seeing streets littered with corpses, women corralled into rape camps, and relatives forced to watch their loved ones murdered or raped.30

The atrocities of Zvornik were just a dress rehearsal for what would befall the Bosniaks over the next three years. The Serb paramilitary groups moved out of Zvornik to the surrounding towns where they would go door to door in search of Bosniaks to humiliate and kill. Soon extermination camps were set up in the countryside to kill Bosniaks more efficiently.31 The Serbs then brought in Bosnian-Serbs from the surrounding areas to inhabit the conquered and “cleansed” towns.32 The Bosnian-Serbs razed the Bosniak mosques and erected Orthodox churches in their place.33 Through their campaign of destruction, the Serbs sought to erase every aspect of the Bosniaks’ presence in Bosnia and Herzegovina.

By the summer of 1992, the Serbs had established a firm grip on all of northwest Bosnia and Herzegovina.34 The Serb forces then took their extermination campaign eastward and southward. Key to the Serb assault on the Bosniaks was the siege of Sarajevo, the capital of Bosnia and Herzegovina. For three and a half years, the Serbs transformed Sarajevo into a living hell.35 The Serb forces implemented an embargo to prevent humanitarian and military assistance from entering the city.36 Serb artillery and snipers followed the embargo by indiscriminately targeting and killing Bosniak civilians in the city.37

The Serb forces dominated the conflict for most of the war; however, the Serbian forces were unable to pull out a decisive victory. The tide of the war began to shift in 1994, when NATO intervened, and the Bosniak forces joined the Croat forces, who had been fighting their own war with Serbia.38 While the war effort began to favor the Bosniak army the longer the conflict went on, the same cannot be said for Bosniak civilians. Faced with a need for a decisive victory to end the war, the Serb forces turned their attention to Srebrenica in western

29. Id. at 114–15.
30. Id.
31. See id. at 115.
32. Id. at 118.
33. Id. at 120.
34. See id. at 122.
35. Id. at 125.
36. Id.
37. Id.
38. See id. at 130.
Bosnia and Herzegovina. The Serbs controlled most of eastern Bosnia and Herzegovina by 1994; however, several cities remained as outposts for displaced Bosniaks in eastern Bosnia and Herzegovina. In 1993, the UN declared several of these outpost cities in eastern Bosnia and Herzegovina to be safe havens directly under UN protection. However, the safe havens proved illusory, as the Serb forces would turn the safe haven of Srebrenica into the site of the most barbaric mass slaughter that Europe saw in the second half of the twentieth century.

The Serb forces instructed the Serb forces to “create an unbearable situation of total insecurity with no hope for further survival or life for the inhabitants of Srebrenica.” The town fell into Serbian hands within five days. One week after the Serb forces took Srebrenica, 25,000 women and children were deported and over 8,000 men and boys were executed in mass graves. The UN safe haven was exposed as nothing more than humanitarian rhetoric. Survivor Vahid Suljić describes hearing the sounds of “Bosniak men screaming and asking for help every night” as the UN peacekeepers stood by doing nothing. As the war increasingly turned in favor of the Bosniaks, the Serb forces would return to the sites of their mass killings, exhume the bodies, chop up the bodies, and then rebury the bodies in other areas of Bosnia and Herzegovina. The denial began before the war even ended.

While a vast majority of the human rights violations in the war were committed by Serb forces against Bosniaks and Bosnian-Croats, the other two forces in the war did not fight with entirely clean hands either. Early in the war, before the Bosnian-Croats joined forces with the Bosniaks, the Bosnian-Croats began their own offensive against the Bosniaks and Bosnian-Serbs. Bosnian-Croats “cleansed” a number of towns in western Bosnia and Herzegovina. The Bosniaks also committed a number of war crimes, as foreign fighters from around the Arab world came into Bosnia and Herzegovina to support their Muslim

39. Id. at 132.
40. Id. at 131–32.
41. Id. at 114.
42. Id. at 131.
43. Id. at 133.
44. ADEM MEHMEDOIĆ ET AL., SREBRENICA MEMORIAL CENTER, SREBRENICA GENOCIDE DENIAL REPORT 2021, 12 (Monica Hanson-Green, ed., 2021).
46. Id. at 127.
47. Id.
These “Mujahadeen” fighters were not regular components of the Bosniak military force and thus were difficult to control, leading to several war crimes committed in the name of the Bosniak cause.

After several years of brutal fighting, the Bosnian war ended with an internationally brokered ceasefire in October of 1995. The warring parties (Bosniaks, Bosnian-Croats, and Bosnian-Serbs) met in Dayton, Ohio, for peace talks which culminated in the Dayton Agreement. The international community crafted a new government structure, which split the country into two entities along ethnic lines. The Federation of Bosnia and Herzegovina represented the Bosniaks and the Bosnian-Croats, whereas the Republika Srpska represented the Bosnian-Serbs. In effect, the agreement rewarded the Serbian aggressors by giving them their own entity within the new country while recognizing the territory they had conquered as legitimate gains.

While the Dayton Agreement may have rewarded the Serbians for their aggression, other international structures looked to hold the perpetrators of the heinous acts responsible. The International Criminal Tribunal for the Former Yugoslavia (ICTY) and the International Court of Justice (ICJ) began prosecuting a number of war actors for war crimes, mass rape, crimes against humanity, and genocide. To this day no court has found the entire Bosnian war to be an act of genocide. Yet, the ICTY and the ICJ have recognized that what happened in Srebrenica constituted Bosnian-Serb genocide against the Bosniaks. Judge Alphons Orie described Srebrenica as one of the most heinous crimes against humanity known to humankind.

The debate over whether or not the Bosnian war, in its entirety, amounts to genocide is ongoing. Bosnian-Serb politicians, Bosnian-Serb academics, and Bosnian-Serb media routinely deny that genocide occurred. The acts of denial take many forms. Chief among Bosnian-Serb denial strategies are active denial, trivialization, acknowledgment of crimes but not genocide, support of perpetrators, refusal to discuss

48. Id. at 128–29.
49. Id.
50. Id.
52. See generally id.
53. Id. at 13, 17.
54. Id. at 14; see Annie Hylton, You Simply Die All Over Again: The Agony of Srebrenica, 25 Years on, THE GUARDIAN, (July 14, 2020, 3:30 AM), https://www.theguardian.com/global-development/2020/jul/14/you-simply-die-all-over-again-the-agony-of-srebrenica-25-years-on.
55. MEHMEĐOVIC ET AL., supra note 44, at 15.
genocide, conspiracy theories, and triumphalism.\textsuperscript{56} Yet, even respected scholars disagree about whether the war in Bosnia was genocide.\textsuperscript{57} Some scholars say the war was clearly genocide, others say it was a tragic civil war, and yet others argue it was an ethnic cleansing.\textsuperscript{58} The argument may seem trivial, but the label matters because, as Brian Bix points out, some labels carry a normative force.\textsuperscript{59} Genocide, perhaps the greatest crime against humanity, certainly carries a normative force behind it in postwar Bosnia and Herzegovina.

\textbf{B. The Law}

Two years ago (2019), on the twenty-fourth anniversary of Srebrenica, High Representative Valentin Inzko said, “I shall advocate the adoption of a genocide denial law by the parliament of Bosnia and Herzegovina. We shall surely have a law next year on the 25th anniversary of the genocide.”\textsuperscript{60} The twenty-fifth anniversary of the Srebrenica genocide came and went; yet no law was enacted. The structure of the government of Bosnia and Herzegovina ensured that the law would never be passed as long as Bosnian-Serbs continued to deny genocide. The government stalemate over the genocide denial law was ultimately only possible due the Dayton Agreement, the very instrument that stopped the genocide.

The Dayton Agreement was an effort by the international community to not only bring the Bosnian war to a close but also figure out how to create a unified Bosnia and Herzegovina state that could exist post war.\textsuperscript{61} The international community believed that creating a government structure that attempted to represent all three ethnicities equally would create a stable society and government.\textsuperscript{62} Dayton did create a government that would represent all three ethnicities, but it also created one of the most complex and inept governance systems in the world. Bosnia and Herzegovina is one state that is comprised of two legal entities, the Federation of Bosnia and Herzegovina and the Republika Srpska as well as the condominium, the Brčko District.\textsuperscript{63}

\begin{thebibliography}{63}
\bibitem{56} Id. at 20.
\bibitem{57} \textit{See} Catić, \textit{supra} note 51, at 3–4.
\bibitem{58} \textit{See} id.
\bibitem{59} BIX \textit{supra} note 11, at 23.
\bibitem{61} \textit{See} Catić, \textit{supra} note 51, at 14.
\bibitem{62} Id. at 2.
\bibitem{63} Ustav Bosne I Hercegovine [C] Dec. 14, 1998 (Revised 2009), art. I § 3 (Bosn. & Herz).
\end{thebibliography}
Federation is further divided into ten cantons and then divided further into 141 municipalities. The state has three presidents, a council of ministers, and a bicameral parliament. The entities each have a prime minister, president, and government. The Federation’s cantons also have a prime minister, government, and parliament. The municipalities then each have their own assemblies and mayors. This structure alone makes passing legislation difficult, but the ethnic quotas at the highest level of government make meaningful reform near impossible.

The state’s tripartite presidency must be composed of a Bosniak, a Bosnian-Croat, and a Bosnian-Serb. The bicameral parliament is similarly broken up into ethnic quotas. The House of Peoples consists of fifteen delegates comprised of five Bosnian-Croats, five Bosniaks, and five Bosnian-Serbs. The House of Representatives consists of forty-two members with twenty-eight members from the Federation and fourteen members from the Republika Srpska. The mandatory ethnic composition of the House of Peoples alone ensures that genocide denial legislation will not pass. Bosnian-Croat and Bosnian-Serb politicians can simply outvote the Bosniaks. Bosnian-Serbs or Bosnian-Croats could even block the vote on the legislation by not showing up to parliament since three delegates from each ethnicity are required for a quorum.

The genocide denial law that High Representative Inzko desired has failed to pass the parliament three times in the last decade. Yet, on July 23, 2021, High Representative Inzko took matters into his own hands by unilaterally enacting the genocide denial law on his own. The law targets two types of genocide denial, denial to incite and triumphalism.

Subsections two and three of the new law are the incitement

---

65. *Id.*
66. *Id.*
67. *Id.*
68. *Id.*
70. *Id.* at art. IV § 1.
71. *Id.* at art. IV § 2.
72. *Id.* at art. IV § 1(b).
73. *See Dzaferovic, supra note 60.*
74. *See Code Criminal art. 145a(2)–(3), (5)–(6) (Bosn. & Herz.). The law in English can be found on the website of the Office of the High Representative, http://www.ohr.int/hrs-decision-on-enacting-the-law-on-amendment-to-the-criminal-code-of-bosnia-and-herzegovina/.*
The law punishes individuals who incite violence or hatred against a group of people, or a member of that group, based on their race, color, religion, descent, or national or ethnic origin. Additionally, individuals who deny, grossly trivialize, or try to justify genocide, crimes against humanity, or a war crime in order to incite violence will also be guilty of incitement. The genocide, crimes against humanity, or war crime must have been established by the International Criminal Tribunal for the former Yugoslavia, the International Criminal Court, a tribunal established by the Charter of the International Military Tribunal, or a Court of Bosnia and Herzegovina. Individuals who violate subsection two or three will be imprisoned for three months to three years, while public officials who violate subsection two will be imprisoned for no less than three years. When the incitement is “likely to disturb public peace and order or which is threatening, abusive, or insulting,” the perpetrator will face imprisonment for at least three years.

The genocide denial law also prohibits genocide triumphalism. Triumphalism is a unique form of genocide denial where perpetrators celebrate the genocide. They do not deny the killings but rather dispute the status of an event as a genocide. Section six of the Bosnian genocide denial law prohibits recognizing, awarding, memorializing, or granting a privilege to a person sentenced for genocide, crimes against humanity, or a war crime. An individual who violates section six faces imprisonment of no less than three years.

While the genocide denial law has been controversial, the unilateral creation of a law in Bosnia and Herzegovina is not uncommon. Previous High Representatives of Bosnia and Herzegovina have created dozens of new laws over the years. The only issue is that the High Representative does not have the power to enact new laws under the constitution of Bosnia and Herzegovina. The High Representative gathers his authority from the Dayton Agreement. Annex X, Article V gives the High Representative the final authority to...
ENSURING DIGNITY FOR THE SURVIVORS AND THE DEAD

interpret the Dayton Agreement.\textsuperscript{84} Notably absent from the Dayton Agreement is the authority of the High Representative to unilaterally promulgate laws in Bosnia and Herzegovina. That power stems from the Peace Implementation Council’s (PIC) Bonn Conclusions of 1998.\textsuperscript{85} Essentially, the High Representative was told by an international coalition to interpret the Dayton Agreement as giving the High Representative the authority to write law. This has drawn criticism as a form of imperialism with the High Representative acting like an old British viceroy.\textsuperscript{86} Despite the criticism, Bosnia and Herzegovina has abided by the High Representative’s laws since 1998.

III. THE BOSNIAN GENOCIDE DENIAL LAW

CRITICISM AND RESPONSE

The Bosnian genocide denial law suffers from two types of criticism. First, the law has been called undemocratic because it was implemented by the Office of the High Representative.\textsuperscript{87} Second, the law limits free speech.\textsuperscript{88} Like most laws that limit freedom of speech, this criticism comes from both a consequentialist and a deontological perspective.\textsuperscript{89} This section will summarize each of these critiques and then respond to each.

A. The Bosnian Genocide Denial Law as an Undemocratic Law

The Office of the High Representative (OHR) has been under heavy criticism since its inception. At face value this institution appears completely incompatible with liberal democracy. The OHR is an international body that receives its power from both the Dayton

\textsuperscript{85} See PEACE IMPLEMENTATION COUNCIL, supra note 12, at art. XI(2).
\textsuperscript{86} See Gerald Knaus & Felix Martin, Lessons from Bosnia and Herzegovina: Travails of The European Raj, 14 J. DEMOCRACY 60, 66-67 (2014).
Agreement and the Peace Implementation Council. The Bosnian people have no direct say in who the High Representative is outside of their representative in the PIC. The OHR has the ability to make binding decisions so long as the decisions seek to ensure the implementation of the Dayton Peace Agreement or the smooth running of Bosnian institutions. Essentially, an unelected non-Bosnian official has the power to craft laws that completely bypass the elected legislature of Bosnia and Herzegovina. This international promulgation of law has led others to see the OHR as a form of liberal imperialism reminiscent of the British East India Company.

Critics from the Republika Srpska claim the OHR's actions are unacceptable because they are undemocratic and limit the autonomy of the Republika Srpska as a represented entity of Bosnia and Herzegovina. The Bosnian genocide denial law High Representative Inzko enacted is not actually a new idea. The Bosnian legislature has attempted to pass a genocide denial law three times before. Each time the law was put before the legislature, it failed to pass. By bypassing the legislature, the OHR is bypassing the democratic process and enacting a law that has failed to pass through a popularly elected legislature three times. The genocide denial law is also an example of the OHR limiting the Republika Srpska's autonomy. Essentially, the OHR has overruled the Republika Srpska's constitutional blocking of the law's passage in the legislature. The Republika Srpska could debate, challenge, or even block a law in the legislature, but when the OHR is involved, the entity has no political means of blocking laws it sees as harmful to Bosnian-Serbs.

While the OHR is not an ideal political institution, nor was it intended to be a long-term fixture of Bosnians governance, the OHR is not an unprecedented institution and is necessary to move Bosnia and Herzegovina closer towards a liberal-democratic society. The international community has often played a role in state-building after intervening in mass atrocities. “Germany, Austria, and Japan all owe their modern democratic regimes to decades of foreign rule and imposition of a sort far less expansive than the OHR.” In response to criticism of the OHR's implementation of the genocide denial law, Emir Suljagić, the head of the Srebrenica Memorial Center, noted that

90. See Gerald Knaus & Felix Martin, supra note 86, at 63–64.
91. See PEACE IMPLEMENTATION COUNCIL, supra note 12, at art. XI(2).
92. See Gerald Knaus & Felix Matin at 61–62.
93. See Memićević, supra note 88.
94. See Dzaferovic, supra note 60.
95. Id.
96. See Mujanovic, supra note 10.
virtually all of the denazification and anti-Holocaust denial process in postwar Germany occurred through an Allied occupation regime.\textsuperscript{97}

The OHR was also not imposed on Bosnia and Herzegovina without consent. The OHR was created by the Dayton Agreement, a treaty that was signed by the Republic of Bosnia and Herzegovina.\textsuperscript{98} Defenders of the OHR argue that by signing the Dayton Agreement, Bosnia and Herzegovina has consented to subject itself to the authority of the OHR, who is responsible “for interpreting the Social Contact which lies at the heart of their [s]tate.”\textsuperscript{99} The OHR’s ability to craft binding laws stems from Annex X, Article V in connection with Article II(1)(D).\textsuperscript{100} Article II(1)(D) instils the OHR with the responsibility to resolve any difficulties with the implementation of the Dayton Agreement.\textsuperscript{101} While Article V gives the OHR the power to interpret the Dayton Agreement for the civilian implementation of the Agreement.\textsuperscript{102} Together these Articles grant the OHR the power to issue binding decisions.\textsuperscript{103} This interpretation of the OHR’s powers was confirmed by the PIC Bonn Conclusions.\textsuperscript{104}

The OHR’s genocide denial law is the type of law that furthers the goals of the Dayton Agreement and ultimately pushes Bosnia and Herzegovina towards a modern liberal-democratic society. Annex VI of the Dayton Agreement deals generally with the security of essential human rights.\textsuperscript{105} This list of rights includes the freedom of expression, but all rights, including freedom of expression, are limited by the understanding that no right shall be upheld to the extent the right allows an individual to discriminate.\textsuperscript{106} The genocide denial law complies with Annex VI’s conception of human rights. The law respects the freedom of speech, for example the law does not criminalize academic debate over whether the status of Bosnia’s war constituted a genocide or an ethnic cleansing. However, the law also seeks to respect the human dignity of the survivors and dead by forbidding the freedom to express hatred that dehumanizes and degrades groups of people.

The fact that the law has failed to pass the Bosnian legislature three times is also indicative that the OHR needed to step in. The

\begin{footnotes}
\footnotetext{97} Id.
\footnotetext{98} See The General Framework, supra note 64, at Annex X.
\footnotetext{100} Id. at 264.
\footnotetext{101} See The General Framework, supra note 64, at Annex X, art. II
\footnotetext{102} See id. at Annex X, Article V.
\footnotetext{103} Banning, supra note 99, at 264.
\footnotetext{104} See PEACE IMPLEMENTATION COUNCIL, supra note 12, at art. XI(2).
\footnotetext{105} See The General Framework, supra note 64, at Annex VI.
\footnotetext{106} Id.
\end{footnotes}
genocide denial law failing three times is not a sign of democratic governance. Rather, the genocide denial law repeatedly failing is indicative of the prevalence of ethno-nationalist politics in the Bosnian legislature. Each time the law has failed to pass, the law was voted down on ethnic party lines led by the Bosnian-Serbs of the Republika Srpska. The OHR was designed to overcome such obstruction or lack of political will to undertake essential reform. As former OHR Lord Paddy Ashdown once said, “[t]he more you reform, the less I will have to. The less you reform, the more I will have to.”

The Bosnian legislature has failed to reform and has allowed denial of horrible crimes against humanity to become a mainstay in Bosnian society. The genocide denial law seeks to rectify this failure. Christoph Heubner, executive vice-president for the International Auschwitz Committee, said the law “would create the conditions for ‘honest and dignified discussion’ about the facts of the genocide in Bosnia.” A frank and open discussion about what happened during the Bosnian war is essential to the peace and reconciliation process in Bosnia and Herzegovina as well as pushing Bosnian and Herzegovina towards a more liberal-democratic society.

B. The Bosnian Genocide Denial Law as Free Speech

Like most laws that limit expression, the Bosnian genocide denial law is susceptible to criticism that the law violates the fundamental right of free speech. This criticism usually stems from a consequentialist and deontological approach.

The most easily recognizable argument against any form of censorship is that speech does not violate the harm principle and is thus outside the moral authority of government to regulate. The poster child for this argument, and classical liberalism generally, is John Stuart Mill. Mill argued “[t]he only purpose for which power can rightfully be exercised over any member of a civilized community, against his will, is to prevent harm to others.” Critics of Mill point out that harm is a rather ambiguous term which fails to offer any concrete categorization

107. See Dzaferovic, supra note 60.
108. See Gerald Knaus & Felix Martin, supra note 86, at 63–64.
109. Id.
110. See Mujanovic, supra note 10.
111. See Schauer, supra note 89, at 28-29.
112. BIX, supra note 11, at 173.
113. JOHN STUART MILL, ON LIBERTY AND THE SUBJECTION OF WOMEN 16 (Alan Ryan ed., 2006).
of what is harm and what is not harm. Yet, Mill’s position is usually qualified to argue that only harmful conduct/physical action should be prohibited by government. This more refined harm principle is called the harmful conduct principle. This clarification avoids some of the most pointed criticisms of Mill’s original harm principle, which assert that a non-trivial harm can be found in every moral offense. This essentially boils every harmful conduct principle issue down to one simple question, is the conduct causing a person physical harm?

Consequentialists would argue the Bosnian genocide denial law exceeds the legitimate government authority because genocide denial and triumphalism do not physically harm anyone. The law does not limit its restriction of denial to only incitement of violence, a position that even most free speech absolutists would not defend, but the law also prohibits denial speech if it will incite hatred toward another group. Simply inciting hatred does not actually harm any individual. People do not have a right to be liked. By its very nature, it could be said that genocide denial practically always incites hatred toward another group. Denial may be hurtful, grotesque, or upsetting; yet, denial is not on the same level as crimes such as robbery, murder, and assault. The words do not physically harm anyone. Therefore, since no real physical harm occurs, the government of Bosnia and Herzegovina has no right or authority to imprison people for speaking unpopular opinions.

From the deontological perspective, the critics admit that speech may harm some individuals, but the right is still a fundamental human right that must trump short term social goods. Organizations such as the American Civil Liberties Union posit that:

“[w]hether it be because allowing harmful speech fosters a desirable attitude of tolerance, or leads to the increased discovery of truth, or respects the process of democratic participation in public deliberation, or manifests the self-expression, self-realization, or autonomy, of speaker or hearer, the values underlying the right to free speech are values not themselves derived from the fact of speech’s harmlessness and are not therefore undercut by the fact of speech’s

114. Bix, supra note 11, at 176.
115. Id. at 174–75.
116. Id. at 174.
117. See Code Criminal art.145a(2)–(3)(Bosn. & Herz.).
harmfulness on particular occasions.”

This approach resembles Ronald Dworkin’s “Rights as Trumps” approach. When a right and the general welfare collide the right must trump the general welfare. This “Rights as Trumps” approach to free speech guards against ill-guided policies that may have negative impacts to liberalism and democracy down the line. The classic example is torture. Even though torture could potentially save lives by preventing a terrorist attack, the rights that are violated when the state uses torture are so fundamental that they should not be trumped for a momentary societal good. However, the “Rights as Trumps” approach is not a “Rights are Absolute” approach. The fundamental right is not absolute and may be overridden when it brushes up against a competing fundamental right.

The Bosnian genocide law would be an unacceptable curbing of free speech and thus a hazard to Bosnia and Herzegovina’s liberal and democratic values. Free speech is necessary to foster the values that make a society a liberal-democratic society. Key to a liberal-democratic society is tolerance, open debate, and truth seeking. The Bosnian war is still recent history and plays a large role in the public memory of Bosnia and Herzegovina. The reconciliation process is still ongoing. In order to successfully reconcile and unify the country, an open discussion must occur. Tolerance of opposing viewpoints, even if the viewpoints are repulsive, is key to this discussion. The open dialogue is crucial to truth seeking and unification. While genocide denial may cause some mental anguish today, protecting freedom of speech is important to the longevity of a liberal-democratic society in Bosnia and Herzegovina.

However, genocide denial does warrant regulation as it is another stage of the genocide itself and causes harm to the dignity of the living and the dead. Genocide scholars typically classify the events of a genocide as belonging to one of ten stages. The final stage of genocide, is not the act, or extermination of a group, but rather denial. In this stage the perpetrators “deny that they committed any crimes, and often blame what happened on the victims.” Denial is an important aspect of genocide because the “[c]omplete annihilation of a people requires the

118. Schauer, supra note 89, at 29.
119. Id. at 30.
120. Bix, supra note 11, at 137.
121. Schauer, supra note 89, at 29, 31.
122. Bix, supra note 11, at 138.
124. See id.
125. See id.
banishment of recollection and suffocation of remembrance. Falsification, deception and half-truths reduce what was to what might have been or perhaps what was not at all.” The denial not only destroys the memory of a group but is also “among the surest indicators of further genocidal massacres.”

The denial stage of genocide is strong in the Republika Srpska. Sixty-six percent of Bosnian-Serbs in the Republika Srpska deny that Bosnian-Serbs committed genocide during the Bosnian war. The denial stage also includes Bosnian-Serb triumphalism with common chants such as “Nož, žica, Srebrenica” (“Knife, Wire, Srebrenica”) and “bit će repriza” (“There will be a repeat”). On January 9, 2022, eight hundred militarized Bosnian-Serb police officers marched through Banja Luka celebrating the Republika Srpska chanting “Srbi ima kao lista, nema više komunista. Za krst, za krst, Nemanjića, za slavu, za slavu, Obilića.” (Serbs are as a list, there are no more communists. For the cross, for the cross, Nemanjića, for glory, for glory, Obilića.) The march was significant because of its triumphalist symbolism and open defiance of Bosnian law. The chant plays on themes of Serbian nationalism and the concept of “Greater Serbia” (returning Serbia’s borders to the heights of the medieval Serbian Empire and ensuring its people are of the Orthodox faith). The Republika Srpska’s police units leading the march were also units that played a prominent role in the atrocities committed during the Bosnian war.

Genocide denial, triumphalism, and secessionist talk in the Republika Srpska all demonstrate why Bosnia and Herzegovina needs the genocide denial law. The final stage of genocide is alive and well in Bosnia. The killings may have stopped, but the assault on Bosniak identity and existence is alive and well. Bosnian-Serb genocide denial seeks to destroy the memory of the dead, to hide their very existence. Triumphalism destroys the dead’s dignity by justifying and glorifying the mass killings. Even in death, the dead continue to be the victims of

127. STANTON, supra note 125.
genocide. The state of Bosnia and Herzegovina was not able to protect the victims, all it can do now is protect their memory and dignity.131

The denial and triumphalism do not only affect the memory of the dead, but they also harm the living, particularly the survivors of the genocide. As Munira Subašić’s quote in the epigraph points out, living in a society where survivors of genocide must live side-by-side with perpetrators is painful enough. Genocide denial harms those who survived by downplaying or obfuscating the horrific events survivors lived through. As an act that is part of a continuing genocide, denial and triumphalism harm the living emotionally by attacking their dignity as human beings and questioning their right to existence. The Bosnian genocide denial law appropriately protects Bosniaks from experiencing the hatred and harm that genocide denial causes. Prohibiting genocide denial in a manner that incites hatred is not appropriate because that hatred is still a form of violence. The hatred that genocide denial intrinsically incites is a continuation of the genocidal campaign against the Bosniaks. Bosnia and Herzegovina cannot let denial persist, and it certainly cannot allow another massacre on the scale of Srebrenica.

Even if we took a deontological perspective, the genocide denial law is still an appropriate curbing of free speech because the freedom of speech in this context is outweighed by the need to preserve human dignity. No country in the world recognizes an absolute right to free speech. In fact, laws against Holocaust denial and hate speech are commonplace in Europe.132 This is because the values that the right to free speech protects are not applicable to all speech.133 For example, racially abusive speech is not truth and does not actually seek truth.134 Denial of the Holocaust is not a search for truth, nor does it deserve tolerance because the existence of the Holocaust falls into “a category of ‘historical facts’ about which denial may be criminalized.”135

Additionally, Holocaust denial, and genocide denial generally, does not promote tolerance in a broad sense. This is because law serves an ideational function. State actions, which include inaction, serve as expressions of societal values and attitudes.136 When the law does not act against such heinous speech as genocide denial, “the law implicitly

132. See Robert Kahn, Holocaust Denial and Hate Speech, in Genocide Denials and the Law 77, 80–82 (Ludovic Hennebel & Thomas Hochmann eds, 2011).
134. See id. at 57.
135. See id. at 54.
teaches that respect for individuals is of little importance.”

Since the values that freedom of speech protect are not always present in every single circumstance, the right must be balanced with other rights in certain contexts. This is because the “Rights as Trumps” approach is not a “Rights are Absolute” approach. The underlying idea behind “Rights as Trumps” is that protecting an individual’s rights, even at the cost of a momentary social good, is beneficial to the rights and values of society in the long run. Achieving this goal requires a balancing of values and competing rights. As Joseph Raz explains, “[t]hough he [the individual] gains from the benefit the right secures to others, the weight and importance of the right depends on its value to those others, and not on the benefit that this in turn secures to the right-holder.”

Freedom of expression is ultimately important to building a liberal-democratic society in Bosnia and Herzegovina, but it must not overtake the right of human dignity. Human dignity has served as the “premier value underlying the last two centuries of moral and political thought.” The state exists to protect individuals, and at a minimum, protecting dignity means ensuring the respect for the intrinsic worth of every person. Indeed human dignity may be the fundamental underlying value that all, or nearly all, human rights seek to protect in some capacity. This is why the concept of human dignity plays such a key role in a number of international texts, such as the Charter of the United Nations; the Universal Declaration of Human Rights; the International Covenant on Economic, Social, and Cultural Rights; and the International Covenant on Civil and Political Rights. “[R]espect for human dignity” is even front and center in the Constitution of Bosnia and Herzegovina and therefore also the core goal of the Dayton Agreement.

By enacting Bosnia and Herzegovina’s genocide denial law, the OHR properly signals that human dignity outweighs the freedom of certain speech that does not promote truth or tolerance. If Bosnia and Herzegovina is to truly protect its citizens from discrimination and protect its citizens’ human dignity, then it must forbid speech that directly attacks or questions an individual’s right to exist. Hannah

---

137. GRAY, supra note 133, at 56.
138. Pildes, supra note 136, at 731.
140. SMITH, supra note 139, at 1.
Arendt claimed that genocide was more than just the physical killing of a group of people but was rather a crime against the human status of the victim.\footnote{143}{See Kahn, supra note 132, at 85.} In essence, it denies them human dignity. Yet, “to respect human dignity is to respect an individual’s standing as being possessed of intrinsic worth.”\footnote{144}{Smith, supra note 139, at 76.} Denial exists as the final stage of genocide because it denies both the dead and the survivors of a genocide human dignity. The denial of the crimes committed during the Bosnian war justifies those crimes and assaults the humanity of the victims. By criminalizing genocide denial and other crimes against humanity, the Bosnian state explicitly affirms the human dignity of the dead and the survivors.

Healing cannot commence until all sides recognize what horrendous actions were done during the war. The law accomplishes this by forcing society to recognize that all sides undoubtedly committed crimes against humanity. The law is not an anti-Bosnian-Serb law. Rather, the law is neutral in its treatment of all ethnicities in Bosnia and Herzegovina by criminalizing any convicted crime against humanity. Without such a law, the ambitions of the Dayton Agreement and the Constitution of Bosnia and Herzegovina to establish a liberal-democratic, pluralist society will not come to fruition.

IV. THE POLITICAL SIDE OF THE BOSNIAN GENOCIDE DENIAL LAW

While it is important to tangle with the legitimate criticism of the law, the international community must also be aware that free speech and democracy are not the value prized by the political elite of the Republika Srpska. The elite, such as Milorad Dodik and Željka Cvijanović, care little about democracy, freedom of speech, and human dignity. Rather, these elites care most about power. Political scientist Jasmin Mujanović dubs these types of elites as “elastic authoritarianism.”\footnote{145}{JašMIN MUJANOVIĆ, HUNGER AND FURY: THE CRISIS OF DEMOCRACY IN THE BALKANS, 2 (2018).} They are a breed of rising far-right nationalist leaders across Europe and have been welcomed into the fold by the most notorious of Europe’s far-right, including Aleksander Vučić (Serbia), Viktor Orbán (Hungary), Marion Le Pen (France), Janez Janša (Slovenia), and Vladimir Putin (Russia).\footnote{146}{See Romeo Franz, EU Must Support Democracy in Bosnia and Herzegovin and Impose Sanctions against Republika Srpska Leadership, Argues Romeo Franz, THE PARLIAMENT (Dec. 14, 2021), https://www.theparliamentmagazine.eu/news/article/a-turning-point-for-bosnia-and-herzegovina; see also Hikmet Karčić, Europe’s Far Right Joins Russia and China Backing Bosnia’s Serb Secessionists, HAARETZ (Jan. 9, 2022),}
These elastic authoritarians consistently mutate their ideological stances to remain in power while purposefully stunting economic and political change for their own good.\textsuperscript{147} The key to elastic authoritarianism is nationalism. Despite the prominent view, popularized by Robert Kaplan, that the Balkans are a hotbed of “ancient ethnic hatreds,” nationalism in the Balkans is a manufactured phenomenon used by local elites to maintain control.\textsuperscript{148} Denialism is a paramount tool in their nationalist toolbox. Refik Hodžić has observed that “there’s a connection between economic robbery and the trauma of the past . . . . Poisonous cynics are making money by undermining our attempts to bring peace. This is hatred as a smokescreen for robbery.”\textsuperscript{149}

At the core of the current dispute over Bosnia and Herzegovina’s genocide denial law is the simple fact that the passage of the law undermines the Republika Srpska elite’s power. They criticize the OHR because it diminishes their own power. They care nothing about freedom of speech, but rather they care that they are potentially losing a time-tested nationalist tool. Every step they take defies the peace in Bosnia and Herzegovina, the Dayton Agreement, and liberalism. The international community, particularly the PIC and EU members, must resist the elite’s call to remove the genocide denial law. The international community must remain strong in its commitment to human dignity, the rule/legitimacy of international law, and the fight against far-right political actors.

V. CONCLUSION

The Bosnian war was an appalling example of nationalist hatred and international inefficiency. Despite the rallying cry “never again,” genocide once again spilt blood on European soil. Twenty-five years later, the peace and reconciliation process is still underway, but it is threatened by false denial and triumphalism of the crimes that occurred during the Bosnian war. The OHR’s attempt to prevent further denial and sanctify human dignity have been met with stern criticism, but that criticism must be met head on by the international community. The law is both legal and necessary to the reconciliation process. The international community, and especially the EU and PIC, cannot be further bullied by far-right elites throughout Europe. Taking a stand to defend Bosnia’s genocide denial law is taking a stand for human dignity

\textsuperscript{147} See MUJANOVIĆ, supra note 145, at 2.
\textsuperscript{148} See id.
\textsuperscript{149} Id. at 18.
and liberal values.