

## TIME, TECHNOLOGY, AND INTERNATIONAL LAW

This panel was convened at 9:00 a.m. on Friday, April 8, 2022, by the ASIL International Law and Technology Interest Group (ILTechIG) in partnership with the European Society of International Law's Interest Group on International Law and Technology. The panel was moderated by ILTechIG Co-Chair, Asaf Lubin of the Indiana University Maurer School of Law, who introduced the speakers: Michal Saliternik of the Netanya Academic College; Christian Djéffal of the Technical University of Munich; and Jenny Domino of the Oversight Board at Meta.

### DISRUPTIVE TECHNOLOGIES AND THE NEED TO FUTURE PROOF INTERNATIONAL LAW

*By Asaf Lubin\**

When one thinks of questions of international law and time, one might naturally gravitate to the intertemporal rule. On the one side is the oft cited dictum of Judge Huber in *Island of Palmas*, where he stated that: “A judicial fact must be ascertained in the light of the law contemporary with it, and not of the law in force at the time such a dispute in regard to it arises or falls to be settled.”<sup>1</sup> On the other side are “living instrument” interpretive approaches, like that put forward by the European Court of Human Rights in *Tyrer v. UK*. In that decision, the Court opined that in its interpretation of the European Convention it would give weight to “present day conditions” and be influenced by “the developments and commonly accepted standards” in general policy.<sup>2</sup> The question of whether to be moored to the past or embrace in the future has fascinated treaty interpretation scholars for generations. But as Rosalyn C. Higgins wrote, “time-related problems in international law are by no means restricted to the intertemporal rule.”<sup>3</sup> Rather, as she put it, “temporal matters are all around us.”<sup>4</sup> From the critical date in territorial acquisition law to the concept of “timely notification” in treaty law and in the law of the use of force, to the doctrine of “reasonable time” in human rights law.

But perhaps the biggest time-related problems are those associated with emerging and disruptive technologies. After all, advancements in the fields of nanotechnology, cybersecurity, spacefaring and space-mining, artificial intelligence, content moderation, robotics, augmented reality, and surveillance, to name but a few examples, are redefining the world around us. These tectonic shifts in technology are dramatically impacting every aspect of international law. They generate a set of normative, institutional, and application-related legal uncertainties that then force us to adopt

\* Indiana University Maurer School of Law.

<sup>1</sup> *Island of Palmas* (Neth. v. U.S.) (1928) 2 RIAA 829, 845.

<sup>2</sup> *Tyrer v. UK*, App. No. 5856/72, Judgment, para. 31 (Eur. Ct. H.R. Apr. 25, 1978).

<sup>3</sup> Rosalyn C. Higgins, *Time and the Law: International Perspectives on an Old Problem*, in *THEMES AND THEORIES: SELECTED ESSAYS, SPEECHES AND WRITINGS IN INTERNATIONAL LAW*, VOL. I, at 892 (Rosalyn C. Higgins ed., 2009).

<sup>4</sup> *Id.* at 875.

new regulatory responses. In this spirit, I posed the following question to the generative AI chatbot, ChatGPT. I asked it, what should be the relationship between international law, time, and fast-changing technology. The chatbot responded: “International lawyers must consider the potential long-term effects of new technologies when developing new rules. They must anticipate the ways in which emerging technologies could change the nature of international relations, and the implications that these changes could have for international law.” But as lawyers how good are we really at anticipating with clairvoyance what might be the effects of future technology so that we can futureproof the law? And more substantively, in what ways might our positivist and formalist rules and doctrines support rather restrict such predictive analysis and prescriptive reform?

Troublingly, “[a]lthough time forms a part of the very bedrock of international law as a legal order and fundamentally determines international law as a discipline, the relationship between time and international law has received only limited attention. To most international lawyers, time appears as simply a technical problem, and mainstream international law doctrine presents international law as essentially atemporal.”<sup>5</sup>

As Co-Chair of the International Law and Technology Interest Group, the panel I am moderating today seek to begin to address this seeming gap in the scholarship. In so doing each of the panelists will offer some preliminary and cross-cutting insights on the way to preserve a rule-of-law that is at once susceptible to both the rule-of-time and the rule-of-technology.

One thing that occurred to me in our discussion is that we especially need to think jurisprudentially about the intersectionality of time and sources in international law. Consider the following set of questions, each could form the basis of a doctoral dissertation:

Treaties, Technology, and Time:

- What is the role of treaty interpretation tools—e.g., Vienna Convention on the Law of Treaties Article 31(3) (“subsequent agreement” “subsequent practice” and other “relevant rules of international law”) and the doctrine of “evolutionary interpretation of treaties”—in futureproofing international law from the effects of fast changing technologies?
- What is the role of treaty suspension tools—e.g., fundamental change of circumstances, impossibility of performance—in futureproofing international law from the effects of fast changing technologies?

Custom, Technology, and Time:

- Which countries should be seen as the “specially affected state” in the formation of customs emerging out of the use of a technology? Is it the actual user of the technology, or the potential user? And what about the victims and potential victims from abuses of the technology?
- Could the “Grotian moment” doctrine serve a helpful tool in responding to the risks of fast evolving technologies?
- Can a country be a “persistent objector” to the introduction of a particular technology or to its effects on its polity?
- Where the particular use of a technology is subject to secrecy, including where the technical features of the technology is protected as trade secret, what could be said about the formation of custom from state practice?

<sup>5</sup> Klara Polackova Van der Ploeg & Luca Pasquet, *The Multifaceted Notion of Time in International Law*, in *INTERNATIONAL LAW AND TIME: NARRATIVES AND TECHNIQUES 1* (Klara Polackova Van der Ploeg, Luca Pasquet & León Castellanos-Jankiewicz eds., 2022).

- Given the increasing role of private corporate actors in the development of technologies, how might their decision making on design and infrastructure impact the evolution of customary rules?

There could easily be similar questions raised about the role of Article 38(1)(c) general principles of law to fill lacunas in the law generated by new areas of practice made possible with technologies, the role of courts and judges in regulating fast changing technology over time, and the place of highly qualified publicists, and their identification, in challenging the power imbalances that technology often produces.

These questions and this broader research agenda are things that our interest group will undoubtedly continue to explore in the years and decades to come.

## TECHNOLOGICAL DEVELOPMENTS AND THE TIME HORIZONS OF INTERNATIONAL LAW

*By Michal Saliternik\* & Sivan Shlomo Agon\*\**

We live in a unique period in the world's history, characterized by an ever accelerating pace of technological development. In diverse areas spanning artificial intelligence, quantum technology, synthetic biology, and outer space exploration, the velocity by which groundbreaking technologies nowadays emerge is unprecedented. Importantly, these technologies have the potential to fundamentally change human life and the natural environment not only for better, but also for worse.<sup>1</sup> They can enable longer, healthier, safer, and easier life for people around the globe while mitigating environmental degradation and natural resource depletion, but they may also be used in manners that facilitate oppression, violence, or inequality, or even present existential threats (think, for example, of advanced autonomous weapons, engineered pathogens, or over-exploitation of outer space resources).

The double-edge sword nature of emerging technologies, combined with the limited maneuvering space entailed by their rapid development, suggests that in order to realize their positive potential and reduce adverse effects, we must adopt and implement appropriate regulatory measures well in advance. Such legal interventions should exist not only at the domestic level, but also at the international one, for technologies carry no passports, and their effects know no borders (and if they do, we might wish to remove them). The question that arises, then, is how to render international law fit for this purpose. Arguably, one of the main obstacles to regulating future technologies lies in the short-term inclination of international law. In what follows, this contribution presents international law's short-term bias and discusses its sources. It then emphasizes the need to extend the time horizons of the discipline longer into the future, with a view to steering technological developments onto a positive course. Finally, it points to some possible avenues for moving in this direction.

### I. THE SHORT-TERM PARADIGM

International legal thought and action tend to be confined by a short-term perspective. Under this temporal paradigm, international legal norms and institutions are devised with a view to

\* Netanya Academic College School of Law.

\*\* Bar-Ilan University Faculty of Law.

<sup>1</sup> See, e.g., WILLIAM MACASKILL, *WHAT WE OWE THE FUTURE* (2022).

solving present-day problems and generating immediate value, while downplaying more distant risks and opportunities. As the UN secretary-general recently put it, international decision-makers are “hobbled by a preference for the present,” often acting as if the “future is someone else’s problem.”<sup>2</sup>

This short-term pattern is prevalent in many areas of international law, most notably those that require accommodation to new technologies. Examples include arms control instruments, which regularly seek to address existing warfare methods rather than tackle evolving technologies *ex ante*; the World Trade Organization rules, which are not adapted to addressing the challenges presented by e-commerce and digital economy, even though their prevalence was anticipated when the rules were negotiated; and the International Health Regulations, which were amended in 2005 in view of the conventional outbreaks of the day, but failed to address emerging threats such as genetically engineered pathogens. In all of these areas, international legal arrangements were designed to address the needs that existed at the time they were formulated, while failing to look ahead even into the problems of the subsequent decade.

Remarkably, even those international legal instruments (mainly soft norms) that explicitly refer to the future usually set forth short timescales that do not stretch beyond a decade or two. The Sendai Framework for Disaster Risk Reduction (2015–2030) provides an illustrative example. While this instrument represents an evolution compared to its predecessor, the Hyogo Framework for Action (2005–2015), in that it covers not only natural but also technological hazards, like its progenitor it is inherently short-termist, projecting only fifteen years ahead. The UN Sustainable Development Goals (SDGs) (2015–2030) similarly look only fifteen years into the future. Moreover, the SDGs’ forward-looking vision is constrained not only by its short time horizon but also by its limited substantive coverage, which “neglects the most extreme threats, such as engineered pandemics or transformative artificial intelligence.”<sup>3</sup>

Admittedly, long-termism is not entirely alien to international law. A long-term regulatory perspective can be traced, for example, in the UN climate change regime, which seeks to address the emerging threat of global warming “for the benefit of present and future generations,”<sup>4</sup> *inter alia*, by setting out long-term climate change mitigation goals.<sup>5</sup> In the technological domain, a long-term orientation can be found in the OECD Recommendation on Artificial Intelligence, which aims to tackle the far-reaching effects of artificial intelligence.<sup>6</sup> Such future-directed international legal instruments, however, are rather limited in their reach and scope, and they represent the exception rather than the rule. In its routine development and operation, international law tends to focus on near-term concerns, often to the neglect of long-term interests.

## II. SOURCES OF SHORT-TERMISM

To be sure, there are many understandable reasons for international law to adopt a short- rather than long-term temporal outlook. Above all, regulating for the distant future raises serious epistemic challenges. No one knows for sure what will happen next year, let alone what will happen fifty years or more from now. This uncertainty raises concerns about using inadequate regulatory tools

<sup>2</sup> UN Secretary-General’s Briefing to the General Assembly on Priorities for 2023 (Feb. 6, 2023), at <https://www.un.org/sg/en/content/sg/speeches/2023-02-06/secretary-generals-briefing-the-general-assembly-priorities-for-2023>.

<sup>3</sup> Maxime Stauffer et al., *Policymaking for the Long-Term Future: Improving Institutional Fit* 17 (2021), at [https://www.legalpriorities.org/research\\_agenda.pdf](https://www.legalpriorities.org/research_agenda.pdf).

<sup>4</sup> UN Framework Convention on Climate Change, 1771 UNTS 107 (May 9, 1992).

<sup>5</sup> Paris Agreement to the UN Framework Convention on Climate Change, Art. 18, 2016 TIAS No. 16-1104 (Dec. 12, 2015).

<sup>6</sup> Recommendation of the Council on Artificial Intelligence, OECD/Legal/0449 (May 22, 2019).

and investing resources in vain on the basis of wrong predictions. Although the problem of predicting the future also exists in the domestic context, in the international sphere, and in particular in the context of technological developments, it is exacerbated by the complexity and magnitude of emerging challenges and the multitude of actors and factors that may affect the way they unfold.

In addition to rational cost-effectiveness calculations, the fact that the future is unfamiliar may give rise to various cognitive biases, which may hinder international legal actors from pursuing long-term regulatory endeavors.<sup>7</sup> One such bias is the availability bias, which suggests that decisionmakers tend to assess the probability of events according to the ease with which such events come to mind, and are more likely to act upon scenarios that are easy to recall or imagine.<sup>8</sup> Since future technological developments are very different from their predecessors, and their nature and effects are often difficult to envision (Which ethical and legal challenges will synthetic biology raise? How will advanced artificial intelligence affect the global job market? What will be the implications of mass-scale outer space mining?), they are less likely to induce legal action than more vivid and concrete threats or opportunities.

However, even if future technological or other developments could be easily anticipated, rational, self-interested international lawmakers might prefer to focus on short-term challenges whose regulation yields immediate gains appreciated by their respective communities, rather than undertake long-term commitments for the benefit of future generations.<sup>9</sup> Yet another obstacle to long-term regulation is presented by the fragmented, decentralized structure of the international legal system. Whereas immediate concerns can facilitate international cooperation and help states and international organizations overcome coordination problems, long-term global challenges provide weaker incentives for cooperation.

Although the tendency of international law to focus on immediate needs can be supported by plausible explanations, it is no longer sustainable. In an increasingly complicated and accelerated world, where technological innovations are rapid and tend to be disruptive rather than incremental, short-termism risks rendering the international legal system obsolete and incapable of effectively handling technological challenges when action is due. It is therefore high time to stretch the time horizons of the international legal system further into the future, and make long-term thinking and action an integral part of the system's operation.

### III. EXTENDING THE TIME HORIZONS OF INTERNATIONAL LAW

In recent years, philosophers and political scholars have shown growing interest in the concept of long-termism, associated with the view that in today's world, more than ever before, we should be particularly concerned with ensuring that the long-term future goes well.<sup>10</sup> Long-termism requires us to regularly think about the future and prepare for it, while allocating the necessary

<sup>7</sup> On the applicability of cognitive biases to international legal actors, see Anne Van Aaken, *Behavioral International Law and Economics* 55 HARV INT'L L.J. 421, 439–49 (2014); Tomer Broude, *Behavioral International Law*, 163 U. PA L. REV. 1099, 1121–30 (2015).

<sup>8</sup> On the availability bias, see, e.g., Amos Tversky & Daniel Kahneman, *Availability: A Heuristic for Judging Frequency and Probability*, 4 COGNITIVE PSYCHOL. 207 (1973); Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1127–28 (1974).

<sup>9</sup> Michael K. MacKenzie, *Institutional Design and Sources of Short-Termism*, in INSTITUTIONS FOR FUTURE GENERATIONS (Iñigo González-Ricoy & Axel Gosseries eds., 2016); Simon Caney, *Political Institutions for the Future: A Five-Fold Package*, in INSTITUTIONS FOR FUTURE GENERATIONS 144 (Iñigo González-Ricoy & Axel Gosseries eds., 2016).

<sup>10</sup> See, e.g., MACASKILL, *supra* note 1; THE LONG VIEW: ESSAYS ON POLICY, PHILANTHROPY, AND THE LONG-TERM FUTURE (Natalie Cargill & John M. Tyler eds., 2021); TOBY ORD, THE PRECIPICE EXISTENTIAL RISK AND THE FUTURE OF HUMANITY (2020); Nick Bostrom, *Existential Risk Prevention as Global Priority*, 4 GLOB. POL'Y 15 (2013).

resources in order to prevent harms and generate benefits. On this view, long-termism calls for intertemporal tradeoffs between maximizing social welfare in the short-term and investing in long-term interests. Yet, at least under its “weak” version, long-termism does not imply that the value of actions taken today is exclusively or primarily determined by their long-term effects. It merely asserts that these effects should be more seriously taken into account.<sup>11</sup> This long-term paradigm rests on the intuitive premises that future people count; that the future, especially in the age of rapid technological changes, is vast in size and could allow for unprecedented value and disvalue; and that there are things we can do to steer the future onto a better course.<sup>12</sup>

Drawing on the burgeoning philosophical literature on long-termism, legal scholars have lately developed the concept of *legal long-termism*, which asserts that legal norms and institutions should play a key role in shaping the long-term future.<sup>13</sup> While acknowledging the epistemic and other difficulties associated with long-term oriented legal interventions, proponents of legal long-termism nonetheless convincingly argue that even under conditions of uncertainty, lawmakers should invest more efforts in preparing for future events. The underlying logic is that the stakes can be so high and the time for response so limited, that even if the probability of certain scenarios is low, it is worth taking early legal measures to mitigate risks and promote positive trajectories.<sup>14</sup>

So far, advocates of legal long-termism have paid little attention to the role of international law in ensuring that the future goes well. In our view, however, international law must become a central site of long-term regulation. As noted earlier, at the beginning of the third millennium, the effects of many events and developments reach beyond national territories. This is especially true for technological innovations that can be easily transferred across state borders. A country can impose strict biosecurity rules to control the production of engineered pathogens in its territory, but if other countries ignore this risk, the effectiveness of national control measures may be limited. Similarly, a country may prohibit the development of certain spyware technologies, but if such technologies are used against its citizens from afar, the domestic prohibition would not protect them. In this global reality, international law must take a leading role in regulating emerging technologies and mitigating their potential future harms.

Greater commitment to long-termism entails several changes in the development and operation of the international legal system. Among other things, it requires incorporation of appropriate learning mechanisms that would allow international lawmakers to reduce uncertainty and be able to better assess pending global challenges. Such enhanced knowledge should be sustained through standardized procedures of data accumulation and analysis, which can be pursued, for example, by scientific advisory boards or other professional bodies. Possessing relevant knowledge, however, is only a first step toward effective regulation of future risks and opportunities. Acting upon this knowledge further requires broad collaboration among states, given the comprehensive and widespread nature of technological and other rising global challenges. Also, as many of these challenges, such as artificial intelligence and synthetic biology, are essentially multisectoral, bearing implications for global economy, security, health, environment, and other domains, their effective regulation may hinge on constructive collaboration among multiple international

<sup>11</sup> Weak long-termism holds that we should be concerned with ensuring that the long-run future goes well, as opposed to strong long-termism, which asserts that in many situations, impact on the long-run future is the most important feature of our actions. See Hilary Greaves, William MacAskill & Elliot Thornley, *The Moral Case for Long-Term Thinking*, in *THE LONG VIEW*, *supra* note 10, at 19.

<sup>12</sup> MACASKILL, *supra* note 1.

<sup>13</sup> Eric Martínez & Christoph Winter, *Experimental Longtermist Jurisprudence*, in *ADVANCES IN EXPERIMENTAL PHILOSOPHY OF LAW* (Stefan Magen & Karolina Prochownik eds., 2023); Christoph Winter et al., *Legal Priorities Research: A Research Agenda 1* (2021), at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3931256](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3931256).

<sup>14</sup> Winter et al., *supra* note 13.

institutions. Sure enough, in its current condition the international legal system may not be so amenable to such collaborations, nor does this system feature other characteristics that may be important for long-term regulation, such as adaptability or effective enforcement. However, notwithstanding these challenges, international lawmakers must acknowledge that the time has come to push the international legal system more decisively toward long-termism, thereby aligning its temporalities with changing global realities.

## INTERNATIONAL LAW BY DESIGN

*By Christian Djéffal\**

Emerging digital technologies pose significant problems at the international level. Issues related to artificial intelligence have been on the international agenda since early on: disinformation and misinformation; lethal autonomous weapons systems; or the consequences of large-scale deployment of large language models such as ChatGPT, Bard, or Luminous. While in many cases, governments, civil society, and even the companies providing the systems have called for international regulation, relatively little has happened. In short, the underlying temporal conundrum for international law is the need for agile and rapid responses to mitigate the long-term risks of technologies. I suggest that the first step is to think ahead: Imagine an international legal framework that enables lawyers and technology developers to shape emerging digital technologies such as artificial intelligence, distributed ledgers, and quantum computing. An international law that helps anticipate the risks of these technologies and mitigate them before they materialize on a global scale. Picture a legal system that guides international legal practitioners to take advantage of the opportunities presented by these technologies and to realize the core values of major international legal treaties such as human rights, the rule of law, and democracy. An international law that allows for innovations that advance the purposes of the international legal order. Such a future version of international law would constantly work toward closer integration of law and technology. What would be the elements of such an international legal order? I suggest that a recent trend in the regulation of technology, i.e., law by design obligation, fits the description we are looking for. Therefore, I will first define these norms and provide examples that already exist in international legal practice, and second, suggest ways in which international legal scholarship and practice could evolve to give effect to international law by design norms.

### I. INTERNATIONAL LAW BY DESIGN OBLIGATIONS

International law by design norms are general obligations to incorporate certain international legal principles throughout the lifecycle of technologies, starting with innovation and development.<sup>1</sup> The idea behind international law by design obligations is to translate legal principles into sociotechnical constellations by obliging those who are creating technologies to accept the respective principles as integral goals of the development process. Instead of regulating in a detailed fashion which requirements and processes have to be observed for a technology to comply with international law, law by design obligations establish overarching objectives that organizations responsible for technology development must achieve. Such obligations create new

\* Technical University of Munich.

<sup>1</sup> Christian Djéffal, *The Normative Potential of the European Rule on Automated Decisions: A New Reading for Art. 22 GDPR*, 81 HEIDELBERG J. INT'L L. 847, 857 (2020).

opportunities to interact with technologies starting with the early phases of innovation throughout their lifecycles.

A specific example of a law by design obligation stems from children's rights.<sup>2</sup> The United Nations Convention on the Rights of a Child (CRD)<sup>3</sup> is a human rights treaty with profound impacts on digital technologies, although it was concluded before the internet was generally accessible. While it was very influential to the progressive stance taken in the United Nations Convention on the Rights of Persons with Disabilities (CRPD), the treaty itself does not mention specific technology issues. Yet, the Committee on the Rights of the Child frequently addressed questions of children's rights and devoted General Comment 25 to children's rights concerning the digital environment.<sup>4</sup> The question of design already lies at the core of the problem description. The Committee stated that "[t]he digital environment was not originally designed for children."<sup>5</sup> This effectively means that there is a challenge to redesign current infrastructures as well as to come up with children's rights by design standards for new technologies. General Comment 25 relies on traditional by design goals such as data protection and privacy by design and cybersecurity by design.<sup>6</sup> However, it also refers to goals that are less established, such as safety by design and universal design.<sup>7</sup> Specific references to human rights by design are made in the context of content recommendations in the context of social media<sup>8</sup> and, in particular, technologies to profile children,<sup>9</sup> games, and digital play.<sup>10</sup> An important aspect is also to design explanations in an age-appropriate manner.<sup>11</sup> The policy measures that are laid out in the convention include instruments that can translate human rights by design requirements into specific technical settings such as "industry codes and design standards."<sup>12</sup> Several more specific measures in that regard are mentioned, which also include obligations with negative goals, such as to "undertake child rights due diligence, in particular, to carry out child rights impact assessments and disclose them to the public, with special consideration given to the differentiated and, at times, severe impacts of the digital environment on children."<sup>13</sup>

## II. CHANGING INTERNATIONAL LEGAL PRACTICE

The example of children's rights has shown the general ways in which international lawyers can put principles into practice. Rather than attempting to regulate technologies, General Comment 25 sets out ways and processes by which the law, lawyers, and children's rights advocates can shape technologies in ways that protect or promote the best interests of children. It is not a specific application of children's rights to technology but rather a set of modes for engagement. This is done to extend the reach of international law and ensure that children's rights can be guaranteed in the digital environment. In this way, international law by design ties in with the theme of this year's

<sup>2</sup> This part builds on Christian Djefal, *Children's Rights by Design and Internet Governance: Revisiting General Comment No. 25 on Children's Rights in Relation to the Digital Environment*, 11 *LAWS* 83 (2021).

<sup>3</sup> Convention on the Rights of the Child 1989, 1577 UNTS 3.

<sup>4</sup> Committee on the Rights of the Child, General Comment No. 25 on Children's Rights in Relation to the Digital Environment, UN Doc. CRC/C/GC/25 (Mar. 2, 2021).

<sup>5</sup> *Id.*, para. 12.

<sup>6</sup> *Id.* (privacy by design is referred to at paras. 70, 77, 88, and 110; cybersecurity by design is referred to at para. 116).

<sup>7</sup> *Id.* (safety by design is referred to at paras. 77, 88, and 110; accessibility by design is referred to at para. 91).

<sup>8</sup> *Id.*, para. 37.

<sup>9</sup> *Id.*, para. 62.

<sup>10</sup> *Id.*, para. 108.

<sup>11</sup> *Id.*, para. 39.

<sup>12</sup> *Id.*, para. 24.

<sup>13</sup> *Id.*, para. 38.



Annual Meeting, “The Reach and Limits of International Law to Solve Today’s Challenges.” International law by design addresses technologies and technological infrastructures over time to enable the goals set by international law. This has important implications for several aspects of international law. Firstly, international law by design norms are an example of a proactive stance in international law. They do not merely focus on remedying violations. They aim to shape the reality in which international law operates in a way such that, at best, violations of the law are prevented. This proactive approach implies that international law must necessarily broaden its scope. It cannot solely consider technologies at a single point in time, but must seek to influence them from the early stages of innovation to the point of use in society. This extension of temporal aspects also means that international legal practice must be ready to adapt to changing circumstances in an interdisciplinary environment. It should be involved in all stages of the development and use of technology to have an impact. This would require people responsible for technology to adapt to innovation and technology development processes such as co-creation or agile software development to ensure that the law can have an impact. International lawyers need to learn how to meaningfully contribute to and intervene in these processes. This is a major challenge for international legal scholarship. Indeed, it is an area in which international lawyers are called upon to develop approaches for adapting international legal practices to new tasks.

In terms of specific approaches, I would like to present the research project Reflective Legal Advice on Innovations and Technology Developments (RechTech), which offers ways in which a proactive international legal practice could be realized. The project aims to explore law as a co-productive force and resource for technology design and to enable collaboration and interaction between law and technology. To this end, the project will develop and refine two intervention approaches targeting processes of innovation and the problematization of human rights in development processes.

The first intervention enables the realization of fundamental principles of international law, including human rights, democracy, and the rule of law. It establishes a co-creation format that combines risk assessment and innovation methods for public administration stakeholders who pursue public goals but normally do not participate in technology development. Using this method, we regularly generate innovative ideas. It is tailored to emerging technologies that are already in use in society, but have the potential to be scaled up. We have tested this approach multiple times in the context of artificial intelligence in workshops providing general information on artificial intelligence and its governance. The workshop begins with an introduction to several AI applications that are already in use. Participants choose an application and go through a guided group exercise with a facilitator that consists of three phases. In the first phase, the group conducts a technology impact assessment of the existing application of the technology. In the second phase, participants begin to co-create new ideas through analogies. They select an idea for the third phase, in which they think about the sociotechnical design of their concept, before reconstructing the values embedded in their suggestions. A concrete example would be to start with AI translation tools also used in public administration, such as the World Intellectual Property Organization translation tool.<sup>14</sup> One idea that came up in several workshops was to build an AI to simplify the language. Reflecting on this idea showed us that international legal values are indeed embedded: such systems promote accessibility for people with disabilities, which is in line with the right to accessibility in Article 9 of the CRPD.<sup>15</sup>

<sup>14</sup> INTERNATIONAL TELECOMMUNICATIONS UNION, UNITED NATIONS ACTIVITIES ON ARTIFICIAL INTELLIGENCE (AI) 247 (2023).

<sup>15</sup> On the scope, see generally Francesco Seatzu, *Art. 9*, in THE UNITED NATIONS CONVENTION ON THE RIGHTS OF PERSONS WITH DISABILITIES: A COMMENTARY (Valentina Della Fina, Rachele Cera & Giuseppe Palmisano eds., 2017).

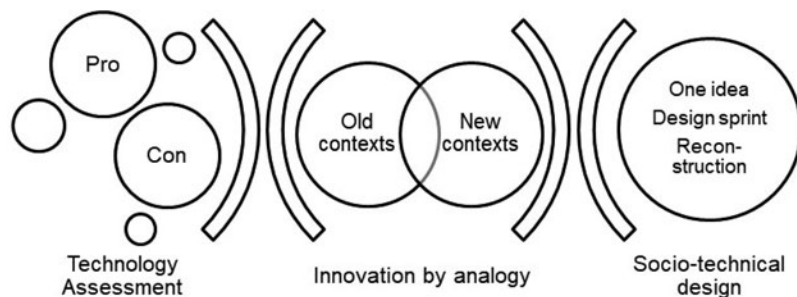


FIGURE 1: RechTech Co-creative Method<sup>16</sup>

The second intervention aims to incorporate material principles of international law like human rights, the rule of law, and democracy within technical infrastructures. It bridges the gap between self-assessment and engagement of those affected by the technologies. The main objective of the second intervention is to confront innovators and developers of robotic technologies at an early stage with the freedoms, rights and interests of those who will use the technologies. We are currently in the process of creating a liability tool that will allow developers, in the first step, to make an initial mapping of the benefits and risks of a given technology. In the second step, they will be guided to present an early prototype to affected stakeholders and learn from this interaction. While on the surface the discussion is about liability law and stakeholder interests, the actual content of the conversation in many cases quickly turns to proportionality tests in the context of human rights.

### III. OUTLOOK

Law and design approaches are becoming increasingly popular in various fields of law. Understanding legal activities more proactively and constructively could open up interesting spaces and avenues for international lawyers. However, this is a challenging process that will require international lawyers to engage with other fields and partially adapt to new environments to effectively communicate international law. The value of these efforts would be to extend the reach of international law in a major societal transformation that is already shaping the infrastructures that also shape the reach and the limits of international law. This makes it not only a fascinating area of research but also a question of what international law and the international community will look like tomorrow.

## HUMAN RIGHTS AND THE CONCEPT OF TIME IN CONTENT MODERATION: META'S OVERSIGHT BOARD

*By Jenny Domino\**

### I. INTRODUCTION

The Oversight Board is now in its third year since it first started accepting cases from Facebook and Instagram users in October 2020. Described then as a “wild new experiment” in online speech

<sup>16</sup> This figure is from the author.

\* Oversight Board at Meta.

governance,<sup>1</sup> this body is tasked to decide the most significant and difficult online speech questions that we see on social media through the lens of international human rights law. The Oversight Board's work evidences the flexibility and utility of international human rights law to shape the long-term future amid rapid technological advances.

How did the Oversight Board come about? In 2018, Mark Zuckerberg announced the creation of a global body that would help ensure accountability from Meta and bring legitimacy to the rules governing the moderation of online content on Facebook and Instagram.<sup>2</sup> The story of Myanmar figured significantly in the development of this body.<sup>3</sup> In 2018, as the Cambridge Analytica scandal was unraveling,<sup>4</sup> so was Meta's inefficient content moderation of anti-Rohingya posts. This was duly reported by the UN Fact-Finding Mission on Myanmar. The UN Fact-Finding Mission found that high-ranking military officials, including the Commander-in-Chief of the Myanmar Army Min Aung Hlaing, other civilian government officials, and ultranationalist monks used Facebook to foment incitement to discrimination and violence against the Rohingya ethnic minority.<sup>5</sup> When Zuckerberg announced the creation of the Oversight Board in late 2018, he mentioned what happened in Myanmar in his blueprint for this governance body, to help prevent a similar set of incidents from occurring again.

Since this announcement, the online speech governance landscape has dramatically changed. The uptick in government regulation, such as the EU's Digital Services Act, have put into question the necessity of self-regulation. This begs the question—is government regulation, or analogously, a treaty, sufficient to prevent harm?

Let us return to Myanmar, this time in 2021. In February of that year, Commander-in-Chief Min Aung Hlaing, whom Meta deplatformed a few years back, staged a coup and arrested political leaders of the National League for Democracy, including State Counsellor Aung San Suu Kyi. As justification for the power grab, the military invoked “voter fraud” in the November 2021 elections, which saw the National League for Democracy's landslide victory. Min Aung Hlaing is now the de facto leader of Myanmar's military regime. In this setting, imagine a government, which by all accounts is not legitimate, enacting regulation. Hardly anyone would contradict that the human rights-informed approach in this scenario is for companies to continue to adhere to the UNGPs, and even resist attempts of governments at “collateral censorship.”<sup>6</sup> The Oversight Board remains a relevant institution in holding Meta accountable especially in countries with governments that cannot be relied upon to protect freedom of expression and other human rights.

## II. ABOUT THE BOARD

The Oversight Board reviews keep-up and takedown decisions made by Meta on Facebook and Instagram. It decides cases based on Meta's content policies, values, and international human rights standards. If a user sees content that they think should be removed for violating Meta's Community Standards, or if a user thinks that a post was mistakenly removed, they can appeal

<sup>1</sup> Casey Newton, *Facebook's New Oversight Board Is a Wild New Experiment in Platform Governance*, THE VERGE (Oct. 23, 2020), at <https://www.theverge.com/2020/10/23/21530524/facebooks-new-oversight-board-platform-governance>.

<sup>2</sup> Meta, *Creating the Oversight Board* (Jan. 19, 2022), at <https://transparency.fb.com/oversight/creation-of-oversight-board>.

<sup>3</sup> Jenny Domino, *Why Facebook's Oversight Board Is Not Diverse Enough*, JUST SECURITY (May 21, 2020), at <https://www.justsecurity.org/70301/why-facebooks-oversight-board-is-not-diverse-enough>.

<sup>4</sup> Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So Far*, N.Y. TIMES (Apr. 4, 2018), at <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>.

<sup>5</sup> UN Human Rights Council, *Report of the Independent International Fact-Finding Mission on Myanmar*, UN Doc. A/HRC/39/64 (Sept. 12, 2018).

<sup>6</sup> Jack M. Balkin, *Free Speech is a Triangle*, 118 COLUM. L. REV. 2011, 2016 (2018).

to the company within the platform to remove or restore the post. Once the user exhausts the appeals process within the platform and the user is still not satisfied with Meta's decision, the user can appeal to the Oversight Board. If the case meets the Board's criteria of significance and difficulty, we select the case for our review. The Board decides whether Meta was correct in deciding to remove the post or whether it should be restored. The Board can also decide, in choosing to restore the post, that a warning screen should cover the post.

The binary keep-up/takedown jurisdiction of the Board leaves out many other types of content: ads, groups, pages, profiles, and events, and other types of decisions, such as content rated "false" by third party fact-checkers.<sup>7</sup> An exception to this would be if Meta itself refers a case or policy advisory request to the Board. The Board was able to weigh in on former President Trump's account suspension in 2021 because Meta referred the matter to us.

Currently, the Board has twenty-two Board Members representing various geographic regions of the world. In October 2022, the Board announced seven strategic priorities that it would focus on: elections and civic space; gender; hate speech; crisis and conflict situations; automation; treating users fairly; and governments' use of Meta's platforms.<sup>8</sup> These strategic priorities were chosen based on an analysis of cases that the Board has received, emblematic of the issues facing many users globally.

### III. SETTING THE STANDARD: INTERNATIONAL HUMAN RIGHTS LAW AS A NORMATIVE FRAMEWORK

As the UN special rapporteur on freedom of expression has stated, human rights law provides a "common vocabulary" to guide content moderation.<sup>9</sup> The Board's work is informed by human rights standards and considerations in three ways: case selection, case decisions, and policy recommendations.

Case selection. The Board selects cases where Meta's platforms have the most adverse human rights impacts. Many of our cases have focused on crisis and conflicts in Ethiopia, Myanmar, Armenia, Azerbaijan, Turkey, and Israel-Palestine. We have also selected cases regarding the elections in Brazil, Cambodia, and the United States, and civic space issues such as the protests in Iran. We have focused on cases where users are treated unfairly, leading to disproportionate impacts on ethnic, religious, and gender minorities, journalists, and human rights defenders.

Case decisions. The Oversight Board can be considered a remedial mechanism under the UN Guiding Principles on Business and Human Rights (UNGPs). In all cases and policy advisory opinions, we consider whether Meta complied with its human rights responsibilities in a particular context in removing or leaving up a post on the platform. In practical terms, we refer to the specific human rights treaty and article provision, General Comment, and relevant UN reports in these decisions. For instance, decisions involving Meta's Violence and Incitement Community Standard have been informed by the UN Rabat Plan of Action's six-factor test to analyze incitement to violence. We conduct a human rights impact assessment to ensure that relevant stakeholders are not exposed to further security and privacy risks.

Case recommendations. The Board issues (non-binding) recommendations that Meta must nonetheless respond to. The Board's policy recommendations try to address the broader systemic issues on how content policies are enforced, such as issues in using automated systems to enforce

<sup>7</sup> Oversight Board By-Laws, Art. 3, Sec. 1.1.2, p. 27.

<sup>8</sup> Oversight Board, *Oversight Board Announces Seven Strategic Priorities* (Oct. 2022), at <https://www.oversightboard.com/news/543066014298093-oversight-board-announces-seven-strategic-priorities>.

<sup>9</sup> UN Human Rights Council, Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, para. 43, UN Doc. A/HRC/38/35 (Apr. 6, 2018).

content policies and automatically close users' appeals.<sup>10</sup> On the surface, we may be looking at an individual piece of content, but our recommendations always try to aim at the broader design and policy gaps that stakeholders have raised as longstanding problems in Meta's products.

#### IV. TIME, TECHNOLOGY, AND INTERNATIONAL LAW

As an institution engaged in applying international human rights standards to social media content, the Oversight Board extends the application of human rights, designed decades ago to prevent state abuses, to present day conditions. Human rights have now become a tool to inform the development of social media company policy and practice. This speaks to the continued utility of the field to shape our future amid rapid technological advances. Already, experts are thinking about how human rights standards can be further applied to other technological developments such as the metaverse<sup>11</sup> or large language models.<sup>12</sup>

A recurring normative question that comes up in our work is how to apply standards originally designed for states in articulating social media companies' human rights responsibilities. Human rights law originated from a state-centric international order, where states were seen as the primary actors in international law. Accordingly, treaties such as the International Covenant on Civil and Political Rights, which enshrines the right to freedom of expression and the prohibition on incitement to violence in Articles 19 and 20, provides specific legitimate purposes that can justify the restriction of expression. As such, companies can ban forms of expression that states are prohibited from regulating under international human rights law, such as spam, bullying content, graphic content, and misinformation. Even the questions posed at the beginning of this panel about the role of treaty interpretation to accommodate emerging technology, or on the formation of custom and the possibility of a "Grotian moment" to respond to new inventions, all speak to states as the primary actors of international law.

Companies differ from states, despite the analogies made about companies as "sovereigns."<sup>13</sup> The consequences of social media platforms moderating user-generated content are different from government regulation of speech by social media users. No one gets sent to jail or fined by companies. There is no deprivation of life or liberty. The worst that could result from persistent content policy violation is an account suspension. For the most part, the violating content gets taken down by the platform.

<sup>10</sup> See e.g., Oversight Board, *Oversight Board Overturns Original Facebook Decision: Case 2020-004-IG-UA* (2021), at <https://oversightboard.com/news/682162975787757-oversight-board-overturns-original-facebook-decision-case-2020-004-ig-ua>; Oversight Board, *Shared Al Jazeera Post*, 2021-009-FB-UA (2021), at <https://oversightboard.com/decision/FB-P93JPX02/>; Oversight Board, *Colombian Police Cartoon*, 2022-004-FB-UA (2022), at <https://oversightboard.com/decision/FB-I964KKM6/>; Oversight Board, *Iran Protest Slogan*, 2022-013-FB-UA (2022), at <https://www.oversightboard.com/decision/FB-ZT6AJS4X/>.

<sup>11</sup> See e.g. Kuzi Charamba, *Beyond the Corporate Responsibility to Respect in the Dawn of a Metaverse*, 30 U. MIAMI INT'L & COMP. L. REV. 110 (2022); Bamali Choudhury, *Business and Human Rights in the Metaverse*, BUS. & HUM. RTS. J. BLOG (Mar. 31, 2023), at <https://www.cambridge.org/core/blog/2023/03/31/business-and-human-rights-in-the-metaverse>.

<sup>12</sup> See e.g., Ido Vock, *ChatGPT Proves that AI Still Has a Racism Problem*, NEW STATESMAN (Dec. 9, 2022), at <https://www.newstatesman.com/quickfire/2022/12/chatgpt-shows-ai-racism-problem>; Kyle Wiggers, *Researchers Discover a Way to Make ChatGPT Consistently Toxic*, TECHCRUNCH (Apr. 12, 2023), at <https://techcrunch.com/2023/04/12/researchers-discover-a-way-to-make-chatgpt-consistently-toxic>; Cass Sunstein, *Artificial Intelligence and the First Amendment* (Apr. 28, 2023), available at [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4431251](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4431251).

<sup>13</sup> See generally REBECCA MACKINNON, CONSENT OF THE NETWORKED: THE WORLDWIDE STRUGGLE FOR INTERNET FREEDOM (2012); Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018).

As the Board's cases on conflict situations grow over time, a related question is to what extent international humanitarian law should apply to platforms when moderating conflict-related posts. This potentially extends the normative reach of the law of armed conflict. In one case, the Board considered whether or not to allow a party to an armed conflict—representing an ethnic minority in the country—to post content that could be interpreted as a threat against the national army.<sup>14</sup> In other cases, the Board addressed the question of violent and graphic content that is shared to raise awareness of human rights abuses.<sup>15</sup> These cases add texture to our understanding of the Geneva Conventions. International law also defines periods of “war” and “peace,” which trigger the application of specific legal frameworks. The reality on the ground, however, is more fluid, and content policies must address the situation on the ground in its state of flux. For instance, elections require policy responses from the pre-election to post-election phase, and incidents of violence may attend these phases to varying degrees, periods that may well fall below the standard of “armed conflict” but arguably not any less consequential.

Time, as a construct under international law, does not provide a straightforward answer on how existing human rights and humanitarian norms should apply in these various cases. Nor does international law provide concrete guidance on how it could apply as various types of technology evolve. Yet, human rights standards nonetheless help ensure that companies do business in a manner that considers the human rights impacts of their operations. This is despite the lack of a treaty that binds companies or of a customary norm whose formation they helped crystallize. The Oversight Board's work, as a global institution applying international human rights standards to novel questions spurred by technology, thus helps in the evolution of international law to respond to technological developments.

<sup>14</sup> Oversight Board, *Oversight Board Upholds Meta's Decision in "Tigray Communication Affairs Bureau" Case (2022-006-FB-MR)* (Oct. 2022), at <https://www.oversightboard.com/news/592325135885870-oversight-board-upholds-meta-s-decision-in-tigray-communication-affairs-bureau-case-2022-006-fb-mr>.

<sup>15</sup> Oversight Board, *Sudan Graphic Video*, 2022-002-FB-MR (2022), at <https://www.oversightboard.com/decision/FB-APONSBVC/>; Oversight Board, *Video After Nigeria Church Attack*, 2022-011-IG-UA (2022), at <https://www.oversightboard.com/decision/IG-OZNR5J1Z>.