global POLICY

GP - ORF Series

Digital Debates

CyFy Journal 2021

Edited by

Trisha Ray and
Rajeswari Pillai Rajagopalan

CyFy 2021
technology · security · society

Durham University

WILEY

Digital Debates

CyFy Journal Volume 08 (2021)

Observer Research Foundation
20 Rouse Avenue, Institutional Area
New Delhi, India 110002
contactus@orfonline.org
www.orfonline.org

ORF provides non-partisan, independent analyses on matters of security, strategy, economy, development, energy and global governance to diverse decision-makers including governments, business communities, academia and civil society. ORF's mandate is to conduct in-depth research, provide inclusive platforms, and invest in tomorrow's thought leaders today.

Editing and Production: Preeti Lourdes John

Design and Layout: Artlab, Chennai

Cover Image Source: Luza Studios/Getty/Royalty-free

ISBN: 978-93-90494-72-9

Citation: Trisha Ray and Rajeswari Pillai Rajagopalan, ed, Digital Debates: CyFy Journal 2021, (New Delhi: ORF and Global Policy Journal, 2021)

Contents

Editors' Note

The most enduring idea from George Orwell's *1984* is that even as we design and produce language, we are shaped in turn by the language we use—words are crucial in explaining complex ideas and in the process of meaning-making. The dynamic and fast-evolving world of technology, innovation, politics, and security sees the equally frenetic adoption of catchphrases. 'Cyber sovereignty,' 'geotech,' 'ethics washing' and more are fast becoming common lexicons, deployed differentially and differently. Sometimes used with care and thought, and on most occasions with an incomplete understanding of the distinct contexts that shape their use.

'New normal' was one such phrase that captured a collective experience, marking an apparent inflection point in our relationship with technology. Yet the term masks the fact that the world we are entering carries many of the same fissures, fractures and fallibilities of the past decade. Or is the encoding of the cleavages of the past in our digital future the 'new normal'?

Conflicts in a Connected World

The past year has brought several complex questions to the fore: Should platforms be able to take decisions that challenge sovereign laws or reinterpret them? Is there a fundamental tension between the global reach of digital giants and the fact that they are primarily shaped by the mores and cultures of small, isolated communities, and subject to the laws, political processes and ethics of the country they are headquartered in? And are global digital norms and values achievable or even desirable?

Our contributors analyse the ongoing 'battles' between the old giant, the many-limbed Leviathan (the State) and the new five-headed giant comprising of technology corporations that shape how we consume, interact and generate value online. The narratives around a 'third way' and propagation of the concept of 'cyber sovereignty' point to a new anxiety that grips nations and communities globally. These drivers are compelling them to autonomously forge a path in the Fourth Industrial Revolution amid tensions between the world's two largest economies, the US and China, and to escape the virtual monopoly of a handful of tech giants.

Governments on their part are 'byting' back. Nikhila Natarajan's essay chronicles the falling out between Capitol Hill and Silicon Valley, one of the few unifying clarion calls that cuts across political lines in the US. Nisha Holla, meanwhile, advocates for countries like India carving their own space in the technology revolution, championing national alternatives to foreign tech giants.

On this note, our digital debates pivot to rules of the road. Principles for digital and emerging technologies are gradually gaining traction across multilateral and multistakeholder forums, yet there remains considerable ground to be covered to transform these disparate statements into unifying courses of action for state and non-state actors alike. The path ahead is not an easy one. While international norms are built upon consensus, the winds of geopolitics have paralysed decision-making at international forums and institutions.

Daniel Porras describes such a state of affairs in space security, with processes at the UN Conference on Disarmament and the UN General Assembly failing to achieve any kind of middle ground. Even as space security talks stall, Victoria Samson adds to the milieu of strategic threats arising from the intersection between space and nuclear. Philip Reiner dives into the history

and current rise in ransomware attacks, emphasising that networked threats require networked solutions, and no one organisation can take on ransomware alone. Nivedita Raju stresses on the urgent need for international governance tools to manage the race to the moon. Smriti Parsheera, using the example of AI norms, describes how, in the absence of achievable consensus, global norms conversations are happening through a variety of non-binding coalitions of the likeminded and soft partnerships. Parsheera asks, however, if these "like-minded" coalitions will become counterproductive, leading to the creation of fragmented frameworks. Arindrajit Basu delves further, in a way questioning the apparent binary between democracies and non-democracies, using the example of foreign intelligence operations to ask whether extraterritorial surveillance by democracies like the US is congruent with the democratic norms they purport to promote.

Similarly in cyberspace, our authors fall along a spectrum when it comes to the conflict between the different ideologies and groups of actors. While Abagail Lawson sees value in cyber norms as tools for managing conflict, James Lewis says that it is high time that norms be translated into repercussions for bad actors. "Reducing the level of conflict," he says, "requires not only norms but consequences for the failure to observe them."

As always, in the great battle of norms, narratives, and behemoths of our digital age, the fact is that it is the everyday lives of communities and individuals that are becoming indelibly linked to technology tools. In the macro world of bits, bytes and algorithms, our contributors caution against losing sight of the human lives at their centre, each with distinct identities, aspirations and needs. As Paul Cadario asks, "Unlike a physical universe operating according to the laws of science, what social and ethical principles would underpin a new, technology-enabled ecosystem to make it sustainable, inclusive, and fair, and keep it that way?"

Even after nearly two years of the COVID-19 pandemic, these interweaving issues remain unresolved. Abhinav Verma's piece on bias in AI in healthcare highlights how the false binaries of gender are force-fit into a binary world of 0's and 1's, leaving individuals with gender identities that fall along the spectrum in the lurch. Samuel Naufeld and Sridhar Ganapathy believe that one solution to better technology-led processes is the creation of an open data ecosystem, but who should lead and fund the creation of this ecosystem? Should it be the State? Should it be a private company, or perhaps a separate 'neutral' entity? If access to these datasets is still price-gated, is it truly 'open' and in the best interest of users?

This year's Digital Debates delve into our uneven path toward several possible tech futures, many of which are marked by inequities along the lines of access, capacity, agenda-setting power, and capital. And while the precise contours of our morrows will always remain unknowable, the essays in this journal and the discussions they feed into at CyFy attempt to critically examine and debate issues at the intersection of technology, security and society, and on occasions bravely offer a suggested approach to adopt.

Trisha Ray and Dr. Rajeswari Pillai Rajagopalan

Harms

Greater Harm Is Not Inevitable: How Global Collaboration Can Reduce Ransomware Threat

Philip James Reiner

Ransomware has rapidly grown in scale and scope. For years, it was a style of cyberattack that was primarily a small-scale economic and criminal threat—even just a nuisance. Ransomware is now much more than just a malicious form of software: it is a thriving criminal industry that poses significant national security, and public health and safety risks. The damage it can inflict to a business is catastrophic, akin to a natural disaster. Multiple factors have combined in recent years to make ransomware the incredibly powerful and impactful form of cybercrime it is today. These include international criminal safe havens, powerful distributed computing capabilities, democratised access to malicious tools, the evolving cryptocurrency ecosystem, and an ever-increasing digital attack surface. As these and other facilitating factors have converged, the solutions needed to address the core challenges have become even more elusive. We are yet again faced with the hard fact that while disruptive technologies have immense benefits for broader society, they also come with unintended consequences. As increasingly networked societies grapple with the ransomware scourge, the criminals behind these attacks also continue to expand their target lists—a threat that will only continue to expand its scope and number of victims around the world.

The factors that have driven the explosion in ransomware attacks have something in common—they are international in nature, not just in terms of the geographic focus of the attacks but also the solutions needed to address them. Historically, any one of these challenges has been almost insurmountable on its own. Take for example the reality of safe havens for cyber criminals. These actors (and their cartels) operate with near impunity because a host country is actively protecting them, lacks the resources and capabilities to stop them, or does not prioritise the issue. Finding effective means for combating the criminal's ability to operate with such freedom has long vexed collaborative international law enforcement efforts (despite some instances of success). The world has come to realise that without greater levels of international collaboration and the establishment of norms and their enforcement, the ransomware challenge will only continue to worsen. It is incredibly difficult to defeat these threats, as they pose unsolved challenges that sit between existing law enforcement regimes and expose the gaps in the limited norms that have been established. International efforts to establish state norms of behaviour will perhaps help inform these initiatives, but the success of these counter-ransomware efforts may well inform ongoing attempts to establish those norms—a mutually reinforcing set of opportunities.

No single organisation or nation can solve the ransomware challenge alone—this is but the latest form of digital extortion that has risen to prominence. The fundamental causes underlying the scourge of ransomware—to include broad underinvestment in basic cybersecurity and hygiene by both industry and government—have long languished without enough attention. But the dynamic is shifting now that such attacks pose a significant threat to critical infrastructure, economic stability and human life; some international collaborative efforts are underway in response. The gravity of this challenge, and the difficulty in making progress against the various elements that make it

so virulent, is what drove the Institute for Security and Technology to establish the international Ransomware Task Force (RTF), a grouping of public and private actors, in January 2021[1]. Arguably, the RTF could have been broader in its focus, and its US-centric recommendations reflect the fact that many of the organisations involved were from that country. The effort, however, set a baseline comprehensive strategy that lays out clear steps that can be taken across nations. This paper will build off that collaborative work to point to where international efforts can make a real difference in reducing the severity of the risks posed by ransomware. This paper provides a brief history of the ransomware threat, discusses the root causes behind the explosion of these attacks in the past couple years, and then—in large part based on the IST RTF report—expands on where international, collaborative efforts can potentially make a difference going forward.

A Brief History of Ransomware

Ransomware is a sub-category of malware, a class of software designed to cause harm to a computer or computer network, used by cybercriminals to render data or systems inaccessible for the purposes of extortion. The US Department of Homeland Security's Critical Infrastructure Security Agency defines ransomware as "an ever-evolving form of malware designed to encrypt files on a device, rendering any files and the systems that rely on them unusable. Malicious actors then demand ransom in exchange for decryption. Ransomware actors often target and threaten to sell or leak exfiltrated data or authentication information if the ransom is not paid[2]." The original ransomware attack is said to have involved sending around floppy disks that would download malicious payloads to seize a system—the "PC Cyborg Virus"—providing a physical mailing address to send your payment to obtain the key and unlock your encrypted systems[3].

Today, ransomware proliferates in diverse ways, including through the exploitation of vulnerabilities and social engineering tactics such as "phishing" emails that deceive employees within an organisation to open attachments that launch the malware that then infects their networks. Once they have penetrated a system, criminals move laterally across networks and encrypt and/or exfiltrate an organisation's data. Ransomware victims are typically prompted with a "ransom note" that informs them that their systems and data have been encrypted, with instructions on how to restore their systems by paying the ransom via cryptocurrency.

Ransomware is a flourishing criminal industry that risks the personal and financial security of individuals, and threatens national security and human life. Victims range from the elderly, small businesses, schools and local municipalities to large governments, hospitals and enterprise companies; nearly every type of institution is regularly targeted, disrupted and held hostage. The problem has intensified in recent years. In 2020, nearly 2,400 governments, healthcare facilities, and schools were victims of ransomware in the US alone[4]. Multiple organisations have assessed the costs of ransomware, and while their exact figures vary, all consistently show a steady increase in the number of attacks—and damaging economic impact. The increased anti-ransomware focus of industry and governments in the past few months may have had a momentary impact on these escalating trends, but overall, the threat continues to expand practically unabated[5]. Most importantly, these trends continue to expand around the globe.

"…while ransomware is a truly global threat, the data indicates that Asia was perhaps the most commonly targeted region. With six nations in the top 10 (Including transcontinental Turkey),

Asia accounted for more than a third (35.70%) of all ransomware submissions in 2020. India had the most ransomware submissions in each quarter this year and was responsible for 14.40% of all submissions in 2020[6]."

Ransomware capabilities have evolved significantly since those early days of floppy disks and snail mail payments, gaining in the sophistication used to encrypt systems, distribute the malicious payload, take advantage of more powerful distributed botnets, and ease of entry via commoditised "ransomware as a service" [7] (RaaS) networks of criminal operators. Perhaps most importantly, the threat has been dramatically fuelled by the increasing ability to collect payments via ubiquitous and anonymous digital payments systems. Unsatisfied with simply freezing up systems for ransom payment, criminals now commonly engage in additional forms of extortion, to include the threat of data-leakage or "double extortion", first seen by Maze in 2019 where they would find and steal sensitive corporate and customer information in addition to encrypting data on systems[8]. Using ever more powerful tools, criminals will continue to refine their extortion tactics. As part of their overall efforts, criminal actors are increasingly able to farm out elements of the "kill chain", to include the collection, research, and analysis of business intelligence to determine the scope of their ransom demands using the "contractors" available through the RaaS model [9].

These techniques are incredibly difficult to address at scale as these criminal groups are aware of how response capabilities work and are shooting the gaps between government agencies and public/private collaborative capabilities. This is further compounded by the reality that many of these criminal actors operate outside the reach of traditional international law enforcement mechanisms. Many ransomware criminals can operate with practical impunity as their countries' governments are unwilling or unable to prosecute this form of crime. In other cases, the organisations executing ransomware attacks may be state-sponsored[10], and may be helping nations evade economic sanctions[11]. For example, in an April 2021 announcement of new sanctions against Russia, the US Department of Treasury made a direct connection between Russia's Federal Security Service (FSB) and ransomware hackers, noting that "to bolster its malicious cyber operations, the FSB cultivates and co-opts criminal hackers, including the previously designated Evil Corp, enabling them to engage in disruptive ransomware attacks and phishing campaigns[12]." Again, these attacks target all domains—healthcare, critical infrastructure, education, municipalities, and businesses large and small. The impact of cyber intrusions to human lives has never been more dire.

International Collaborative Efforts Currently Underway

It is important to note there is an increasing international collaborative focus on the ransomware problem. In June 2020, G7 leaders demanded greater accountability and action from Russia in the collective fight against ransomware[13]. Additionally, in October 2020, the G7 finance ministers issued a joint statement calling upon nations to implement the Financial Action Task Force standards to reduce ransomware and other cybercrime[14]. At the same time, the World Economic Forum has brought together its powerful Partnership Against Cybercrime grouping to work on the ransomware challenge and drive actionable solutions[15]. Finally, in June 2021, the US-EU Ministerial Meeting on Justice and Home Affairs included the launch of a US-EU joint working group on prevention and enhanced law enforcement cooperation to address the rise of ransomware attacks in both regions[16]. Again, too much of this remains focused on the US and the

EU. As the ransomware criminal continues to expand their geographic areas of focus, the threat will evolve past these areas of responsibility.

Nations that may not currently be facing significant threats from ransomware gangs must consider—while they still can—steps to prepare and raise the bar that these criminals are forced to overcome. Nations and businesses around the world continue to underinvest in basic cyber hygiene and the tools that will protect them against ransomware and other potent cybersecurity threats[17]. Long-standing efforts to address cybercrime more broadly, and the ransomware specifically, include No More Ransom[18], the Europol-led Joint Cybercrime Action Taskforce[19], and related successful international takedown attempts as was seen against the Emotet botnet in 2020-21. Emotet is a botnet known to support the distribution of the Ryuk ransomware, and the takedown effort involved a worldwide coalition of law enforcement agencies across the US, Canada, the UK, the Netherlands, Germany, France, Lithuania, and Ukraine. Together they disrupted and took over Emotet's infrastructure located in more than 90 countries[20].

Opportunities for Expanded International Collaboration

Much more can be done at the global collaborative level via the diplomatic, economic/financial, operational, law enforcement, technical, and educational lines of effort to slow the pace of ransomware threats and reduce the monetary incentive to engage in these attacks. Any collaborative, multistakeholder approach to countering cybercrime, including ransomware, must be nimble and be able to function at scale. Much of the substance of the following recommendations comes from the IST Ransomware Task Force, and must be executed in tandem and not as standalone. These efforts will require close coordination between governments on the multinational stage and departments and agencies within a country, and close collaboration between industry, governments, and civil society on the local, national, and international levels. These recommendations should be understood to be top-line concepts that will need to be fleshed out in terms of who must take ownership and responsibility for initiation and follow-through at each level:

Diplomatic: Coordinated, international diplomatic efforts must proactively prioritise ransomware through comprehensive, resourced strategies, including using a carrot-and-stick approach to direct nation-states away from providing safe-havens to ransomware criminals. National-level leaders must put this on the agenda, at every opportunity, and speak to solutions and the consequences for inaction. Those discussions must be accompanied by action."Ransomware as a Service": there are different "partners" that are responsible for different functions/phases in the process. As an example, one will build the encryptors, and another the victim shaming websites. Another will engage in the initial network attack, while others are responsible for moving laterally within a compromised network, stealing data, and deploying the ransomware payload. The extortion payments are divided up among these actors.

Economic/financial: A lack of action taken to reduce ransomware actor behaviour within a given nation's borders should have consequences, such as economic and trade sanctions; constraint on "safe-haven" country activity in international financial markets; using evidence of complicity to "name and shame" those actors and their abettors in public forums and disrupt their freedom

of activity; withholding of military or foreign assistance aid; or denying visas to citizens who seek to travel to other nations. This will not be effective without international agreement that these steps are necessary and effective. Additionally, international financial collaboration is overdue to address gaps in the evolving cryptocurrency ecosystem. The cryptocurrency sector that enables ransomware crime[21] needs to be more closely regulated, which will only succeed if approached collaboratively around the globe. Governments must require cryptocurrency exchanges, crypto kiosks, and over the counter trading "desks" to comply with existing regulations, including 'know your customer', 'anti-money laundering', and 'combatting financing of terrorism' laws. Finally, governments and international institutions should consider establishing cyber response and recovery funds to support ransomware response and other cybersecurity capacity.

Operational disruption: The notion of deterrence is often belittled in the cybersecurity domain. When considered relative to the threats posed by ransomware, however, the criminals involved must be deterred through a combination of imposed costs—including the need to disrupt their ability to deploy and maintain their technical infrastructure and payment distribution systems. As these criminals take advantage of "jurisdictional arbitrage"—playing different law enforcement regimes off one another and shooting the bureaucratic gaps to get away with their crimes—international collaborative actions driven through concerted, resourced, sustained operational disruption campaigns targeting chokepoints, affiliate networks, intermediaries, and the ransomware actors themselves must be more effectively shored up and executed. More traditional enforcement mechanisms such as those employed in the Emotet takedown are but a critical piece of global cybersecurity. Efforts against the ransomware threat must also focus on the more immediate "takedown" or disruption of infrastructure, which more strategically aligns with the needs and priorities of many victims and is a significant public interest.

Law enforcement: Internationally coordinated law enforcement efforts have seen some success against ransomware actors, such as in the takedown of Emotet. These can, however, be dramatically expanded, enhanced, resourced, and empowered, particularly with an eye towards the future when ransomware is highly likely to not have the preponderance of its victims in the US and the EU.

Broader access to applicable resources: The reality remains that cybercrime flourishes due to historical underinvestment in cyber hygiene. A prominent element of focus in of the RTF was the need for access to resources. Expanded access through awareness campaigns that highlight basic tools that are readily available—and often free—to help prevent a ransomware attack or limit the scope of damage in the event of an attack could greatly contribute to baseline cyber hygiene steps that would make it more costly for these ransomware actors to engage in these crimes. International efforts, likely through non-profit organisations, to identify and distribute these tools could create a new focus area for global collaborative action against cybercrime. An internationally coordinated effort should develop a clear, accessible, and broadly adopted framework to help organisations prepare for, and respond to, ransomware attacks. In some under-resourced and more critical sectors, incentives (such as fine relief and funding) or regulation may be required to drive adoption.

Expanded education: As part of efforts to make it harder for criminals to break into systems, education for preparation and response to ransomware activities should be expanded. The importance of this for collaborative international efforts cannot be overstated. While one of the

simplest and likely cheapest means for combating the ransomware threat, it is often overlooked in broader strategies to stem the tide. With an appropriate level of emphasis, and resources, basic education on how to prevent and protect against ransomware attacks—particularly in nations that are not yet hard hit—will go a long way to raising the bar and increasing the costs for ransomware attackers.

Conclusion

The actions detailed in the RTF report need to be enacted together and must be coordinated at both a national and international level. If the proposed framework is fully implemented, the international community could see a decrease in the volume of these attacks within a year. With every recommendation, the RTF members worked through the practical implications, and in most cases presented immediately actionable ideas. Ransomware has become too large of a threat for any one entity to address, and the scale and magnitude of this challenge urgently demands coordinated global action—no one can do this on their own.

The ransomware challenge straddles the national security and technology communities. The ability to translate between both public and private entities across domains through deep, trusted interactions is needed for creative solutions and the ability to work directly with both federal leaders and industry partners on the implementation of necessary actions. Neutral parties, such as non-profit organisations, are uniquely positioned to facilitate communication and cooperation between the government and the private sector in our common interest to collectively defend against ransomware attacks.

No single organisation can take on ransomware alone. Neither can any one nation. It is an international problem that requires international solutions to prevent greater harm.

Endnotes

1. Ransomware Task Force, *Combating Ransomware: A Comprehensive Framework for Action*, Institute for Security and Technology, 2021, *https://securityandtechnology.org/ransomwaretaskforce/*.

2. Cybersecurity & Infrastructure Security Agency (CISA), "Ransomware Guidance and Resources," Cybersecurity & Infrastructure Security Agency (CISA), https://www.cisa.gov/ransomware.

3. Crowdstrike, "History of Ransomware," *Crowdstrike*, https://www.crowdstrike.com/cybersecurity-101/ransomware/history-of-ransomware/.

4. Emsisoft Blog, "The State of Ransomware in the US: Report and Statistics 2020," *Emsisoft*, January 18th, 2021, https://blog.emsisoft.com/en/37314/the-state-of-ransomware-in-the-us-report-and-statistics-2020/.

5. Coveware Blog, "Q2 Ransom Payment Amounts Decline as Ransomware becomes a National Security Priority," *Coveware*, July 23, 2021, https://www.coveware.com/blog/2021/7/23/q2-ransom-payment-amounts-decline-as-ransomware-becomes-a-national-security-priority.

6. Emsisoft Blog, "The State of Ransomware in the US: Report and Statistics 2020." An important caveat, most of the submissions Emsisoft tracks are automated single machine ransomware campaigns. Roughly 89% of their reports are the type that most businesses don't care about: STOP (Djvu): 71.20%, Phobos: 8.90%, and Dharma (. cezar Family): 7.90%. Thank you to Allan Liska from Recorded Future for highlighting this data point in the research.

7. "Ransomware as a Service": there are different "partners" that are responsible for different functions/phases in the process. As an example, one will build the encryptors, and another the victim shaming websites. Another will engage in the initial network attack, while others are responsible for moving laterally within a compromised network, stealing data, and deploying the ransomware payload. The extortion payments are divided up among these actors.

8. Jeremy Kennelly, et. al., "Navigating the MAZE: Tactics, Techniques and Procedures Associated With MAZE Ransomware Incidents," *FireEye Threat Research Blog*, May 7, 2020, https://www.fireeye.com/blog/threat-research/2020/05/tactics-techniques-procedures-associated-with-maze-ransomware-incidents.html.

9. U.S. Congress, House of Representatives, Committee on Energy and Commerce, *Stopping Digital Thieves: The Growing Threat of Ransomware*, 117th Congress, 1st Session, 2021. https://docs.house.gov/Committee/Calendar/ByEvent.aspx?EventID=113927

10. Patrick Howard O'Neill, "Inside the FBI, Russia, and Ukraine's failed cybercrime investigation," *MIT Technology Review*, July 8, 2021. https://www.technologyreview.com/2021/07/08/1027999/fbi-russia-ukraine-cybercrime-investigation-ransomware/.

11. United States Department of the Treasury, *DPRK Cyber Threat Advisory: Guidance on the North Korean Cyber Threat*, (Washington DC: Department of the Treasury, 2020), https://home.treasury.gov/system/files/126/dprk_cyber_threat_advisory_20200415.pdf.

12. Department of the Treasury, United States Government, https://home.treasury.gov/news/press-releases/jy0127.

13. White House, *Carbis Bay G7 Summit Communique (*Carbis Bay, 2021), https://www.whitehouse.gov/briefing-room/statements-releases/2021/06/13/carbis-bay-g7-summit-communique/.

14. Abdullah Khan, "G7 finance ministers urge countries to adopt FATF standards against cybercrime," *S&P Global Market Intelligence*, Oct.13, 2020, https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/g7-finance-ministers-urge-countries-to-adopt-fatf-standards-against-cybercrime-60717238.

15. World Economic Forum "Partnership on Cybercrime" World Economic Forum, https://www.weforum.org/projects/partnership-against-cybercime.

16. Kimberly Underwood, "U.S. and EU To Collaborate Against Ransomware." *The Cyber Edge*, 24 June 2021, https://www.afcea.org/content/us-and-eu-collaborate-against-ransomware.

17. Gareth Pereira, "Cybersecurity in ASEAN: an urgent call to action." *Kearney*, 2018. https://www.kearney.com/digital/article?/a/cybersecurity-in-asean. "Because of these factors,

the top 1,000 ASEAN companies could lose $750 billion in market capitalization, and cybersecurity concerns could derail the region's digital innovation agenda—a central pillar for its success in the digital economy."

18. The main backers of the No More Ransom effort are Europol, the Dutch Government, McAfee and Kaspersky. https://www.nomoreransom.org/en/index.html.

19. Europol, "Joint Cybercrime Action Taskforce (J-CAT)," Europol, https://www.europol.europa.eu/activities-services/services-support/joint-cybercrime-action-taskforce.

20. Andy Greenberg, "Cops Disrupt Emotet, the Internet's 'Most Dangerous Malware'," *WIRED*, January 27, 2021, https://www.wired.com/story/emotet-botnet-takedown/.

21. United States Department of Justice, Office of the Deputy Attorney General, *Cryptocurrency Enforcement Framework*, (Washington D.C, 2020), https://www.justice.gov/archives/ag/page/file/1326061/download.

22. "Cryptocurrencies can vary in their degree of anonymity depending on the public or non-public nature of their associated blockchain. For instance, while Bitcoin addresses do not have names or specific customer information attached to them, Bitcoin's blockchain is public. As a result, users can query addresses to view and understand Bitcoin transactions to some extent. Other cryptocurrencies, however, use non-public or private blockchains that make it more difficult to trace or to attribute transactions. These are often referred to as "anonymity enhanced cryptocurrencies" ("AECs") or "privacy coins." Examples of AECs include Monero, Zcash, and Dash." Additionally, on the challenges facing the cryptocurrency ecosystem, see: Nikilesh De, "State of Crypto: Ransomware Is a Crypto Problem," *Coindesk Opinion*, June 8, 2021, https://www.coindesk.com/policy/2021/06/08/state-of-crypto-ransomware-is-a-crypto-problem/.

Gender as a 0-1 Binary: Future of Gender and Sexual Identity in India's Digital Health Journey

Abhinav Verma

The COVID-19 pandemic has caused a paradigm shift in healthcare delivery, integrating and relying heavily on digital tools to patch over gaps in the overburdened physical health infrastructure. However, the digital transformation of India's healthcare had commenced in 2015 with the 'Digital India' campaign and revitalised through the National Digital Health Mission (NDHM) in 2020. This policy push has propelled the Indian health-tech industry—the country had approximately 4,892 startups as of 2020, and artificial intelligence (AI) in the health space is predicted to grow to about US$6 million by 2021[1].

Digital health initiatives could potentially improve couple communication by taking health discussions outside the realms traditionally held by women-only clinical settings, and also enhance women's decision-making, social status and access to health resources, by establishing new modalities of accessing health information and advice outside of spousal or familial control[2]. During the pandemic, sexual and reproductive health services have been delivered online, at times surpassing issues of equity, access and affordability, especially in certain remote settings[3]. Improved and cost-efficient health service delivery hinges on the creation of longitudinal patient records through a unique identifier, a digitised service delivery mechanism (such as through tele-medicine), central data repositories for further processing and population health analyses, and digital health tools that rely on AI systems to enhance human capabilities in service delivery. Nevertheless, much remains to be done to ensure that existing biases in the health system, especially those related to gender, are not mirrored and amplified in digital systems.

The algorithmisation of healthcare and health service delivery is a double-edged sword—neglecting sex and gender differences in health data and AI development within health tools will lead to suboptimal care for certain groups and active discrimination, unless certain foundational reforms are instituted at the outset. This is more pronounced for individuals that do not fall within the arbitrary binary construct of biological sex. Evidence shows that individuals who do not match the common criterion for assignment to either sex are among the most marginalised and underserved populations in global health systems[4]. This is especially true in India where social and structural barriers to health programmes exist and awareness about healthcare rights is lacking[5].

Gender as the Input

As digital records in health systems become ubiquitous, the delivery of health services to an individual at the point of treatment becomes intrinsically linked to the data stored therein. Furthermore, these records, especially electronic health records (EHRs), have created large repositories of data suitable for population level analysis and further advanced data analytics[6]. AI can now examine, develop, and predict based on this information. This type of electronic phenotyping seeks to discover patient characteristics beyond the input data present in EHRs

using large statistical analysis, for example to sniff out undiagnosed diabetes[7] or even the risk of mental illness and suicid[8]. Biological sex is one of the many input variables that determine these predictions.

Globally, there is increasing evidence indicating that comprehensively capturing gender and sexual orientation identity data can lead to a better provision of care, reduced harm of misdiagnosis, and early and adequate provision of lifesaving care. It can also present pivotal analytical details at a regional or national level for identifying and reducing disparities. In 2013, the UC Davis Health became the first US academic health centre to incorporate this demographic into an EHR, and found that it propelled culturally competent care for LGBTQIA+ populations[9]. It also allowed doctors to take decisions to enable the better provision of preventive care among these populations, including more frequent pap smears for lesbian patients with a higher risk for cervical cancer or the early administration of vaccinations against the papilloma virus among young LGBTQIA+ patients[10].

Misgendering in Electronic Health Records

Traditionally, EHR systems across the world have over-relied on biological sex as the primary and binary attribute to understand gender identity, usually due to a lack of comprehension of the differences between sex and gender. Most EHR systems use either sex or gender as the field to determine what is basically just the biological sex. In India, the 2016 EHR Standards place reliance on the 2011 Data and Metadata Standards (MDDS) for the purpose of demographic detailing[11]. MDDS does not feature a field for "sex" but only "gender", even though the options available—male, female, other—seem to be derived from a biological essentialist understanding of gender. Since MDDS are commonly used across all domain applications for semantic interoperability among e-governance applications, one can assume that the gender-sex confusion permeates across many other digital public goods in the country.

This lack of delineation between sex and gender in EHRs further perpetuates the inequities experienced by the transgender and gender diverse (TGD) populations. EHRs do not account for collection of structured sexual orientation data, making individuals who are excluded due to the gender binary invisible to their health service providers. This impedes provision of important healthcare services for LGBTQIA+ individuals, such as appropriate preventive screenings, assessments of risk for sexually transmitted infections and HIV, and effective intervention for behavioral health concerns that may be related to experiences of anti-LGBT stigma[12]. More importantly, anatomical inventory and gender identity integrations into EHR could enable gender-affirming care for TGD patients[13].

In a traditional EHR system, misgendering may mean missing out on the prostate cancer risks that transwomen and non-binary individuals carry because they may be listed as female without an EHR capturing their comprehensive anatomical inventory. Besides this, giving doctors better visibility over the identity and preferences of their patients can help them create a more welcoming and affirming environment for the provision of care, for instance by using appropriate pronouns and setting-up sessions with counsellors if necessary. A comprehensive and inclusive EHR can enable retrospective studies to detect disparities in services to LGBTQIA+ patients, thereby making appropriate policies for redressal[14]. In the absence of a patient-level data source to this effect, evidence that can potentially affect positive change simply does not exist.

Bias Amplification in Other Input Datasets

Latest developments in state policy indicate a commitment to ensuring data availability for public and private research and the development of digital health tools, primarily focused on stratifications, classifications and predictions. The 2020 National Health Data Management Policy under the NDHM has also brought to light the possibility of enabling access to anonymised and de-identified personal health data (whether contained in EHR or otherwise) to companies and other entities for the purposes of health and clinical research, archiving, statistical analysis, policy formulation, and the development and promotion of diagnostic solutions. In July 2021, Department of Biotechnology released the Biotech-PRIDE (Promotion of Research and Innovation through Data Exchange) guidelines to enable the sharing and exchange of high-throughput, high-volume biological data generated by research groups across the country.

Outside of misgendering, there are other implicit biases in health datasets that can be further amplified across the value chain, including the collection of biased data during first encounters with the health system and researchers, the integration of these biases into algorithmic development, and finally reflecting in the outputs of the analytical process (usually a result of a black box that makes reverse tracking of the decision-making matrix a near impossibility)[15]. Many clinical studies or trial datasets across the world have had a systemic neglect of sex-specific biological differences and almost no representation from TGD populations[16]. Thus, bias is not created by the machine, but only more pronounced by automation. For instance, epidemiological research shows a greater prevalence of depression among women, perhaps since clinical scales of depression study symptoms that occur more frequently among women[17].

Digitisation has quickly enabled novel forms of data to be informative of patient health, even sometimes substituting the traditional process of physical interrogation of a patient's body. Internet of Things-enabled products and services (such as fitness trackers, social media footprints and retail purchases) now have computational value as Big Data towards both population and individual health analyses. These are called digital biomarkers, many of which are currently being presented before regulators like the US Food and Drug Administration (FDA) with clinical validation studies for use in risk detection, diagnosis and monitoring of maladies[18]. The same issues of underrepresentation with clinical trials for drug discovery are also present in trials to show efficacy of digital biomarkers for specific use-cases. They also show insufficient demographic information on sex and gender identities[19]. For example, research shows that digital biomarkers obtained from patients' personal smartphones could be a meaningful modality to capture real-time, in vivo, longitudinal behaviours and self-assessments on mental health[20]. But in countries like India, where women are reported to be 28 percent less likely to own a mobile phone and 56 percent less likely to use mobile internet as compared to men[21], such a model will inevitably be based on uneven samples, thereby promoting misrepresentation and automation risks to its consumers, unless proper safeguards in the regulatory approval pathways are instituted.

Gender as the Output

There is a global wave of AI-based applications with a primary purpose to identify or recognise gender. Big Tech companies like Microsoft, Amazon and IBM were at some point investing in computer vision technologies to recognise binary labels of biological sex[22], even though the

technologies are clubbed under "automated gender recognition" (AGR) that can go beyond computer vision to include speech or facial recognition and even social media scrubbing for classifying the biological sex. There could be multiple different use-cases for such technologies. Companies like Amazon might want to classify shoppers according to gender to make future inventory decisions. However, there is a legitimate concern that these technologies might fall victim to real-world applications that enable more active discrimination[23].

The issue with these technologies is that algorithms have no role to play in taking decisions about another's gender identity, purely because gender identity is determined by the autonomy of the individual. Bringing biological essentialism into the realm of technology only overturns decades of progress made in accepting self-identity as the core of gender recognition. Because AI is trained on data that is put together with majority cis-binaries and labelled by third parties making decisions about classification (essentially inferring gender on a picture without confirming with the individual in the picture), it is natural that such algorithms will churn out binary outputs.

Research on AGR shows that these fears are not unfounded. A content analysis on 58 leading AGR publications shows that 94.8 percent of the time, papers treated gender as binary; 72.4 percent of the time, they operated on an implicit assumption of gender being immutable; and in nearly 60 percent of instances, used physiological characteristics to classify gender[24]. Reports suggest that some trans Uber drivers in the US had their accounts suspended and lost their jobs due to the use of a facial recognition-based security feature that was ill-trained at identifying the faces of individuals who were transitioning[25]. If such technologies are used in the public sphere, particularly for services or spaces that are inherently gendered like bathrooms, it can lead to machine-induced inaccessibility for TGD populations.

The integration of these technologies into health settings could have severe ramifications for TGD populations when viewed in their biological or sociological binaries without accounting for their unique medical and anatomical history, ranging from denial of service altogether (for instance, transwomen who may not have transitioned are refused prostate check-ups), incorrect prognosis (for instance, transmen not being referred to cardiovascular emergencies swiftly enough even though it is known that the risk is higher among men), and even wrongful procedures (such as a pap smear test on a transman who does not retain a cervix after gender-affirming surgery). Automation simply cannot account for the complexities of medical care across the gender spectrum. AGR is also being experimented with across the world in public health, for example with the usage of voice patterns to detect vocal fold cysts[26]. What further complicates this issue is the lack of enthusiasm among researchers to make special and affirmative actions towards including the gender spectrum[27].

Any widespread application of AGR, either through private innovators or within the public system, will inevitably lead to collapsing the self-identification principle of gender identity and police gender expression. It is clear why medical professionals and innovators see immense promise in these technologies. Since most existing medical training and processes are designed only for situations where someone's gender consistently matches their anatomy and resulting medical needs[28], AGR only helps create more certainty and standardisation for healthcare providers. But it cannot be overstated that such applications will also consequently amplify accessibility issues that have always existed within health systems for TGD populations[29], and as most research studies claim, the small size of the population that faces gender dysphoria might be too insignificant to consider in the face of blind technology solutionism.

Mitigating Bias in Digital Health

In an ideal scenario, design and human-computer interface considerations should be accounted for before the product goes live. As India matures across its digital health journey, it is important to identify low-hanging fruits (or saving graces) to safeguard the most critical interests of the TGD communities and citizenry at large against biases that might creep in, especially focusing on those measures that have multiplier impacts across the algorithmic value chain. There are many pivotal sociological interventions that are missing in the health system and that can make it more gender-friendly than the pure technology-based corrections proposed below. Technology solutionism is not the most critical redressal; in fact, it is far from it. However, this paper's focus is limited to viewing the issue from data, technology and the larger 'infrastructuring' lens, and the recommendations are made within its ambit.

Immediate reform for EHRs as a product and process

As greater exploitation takes place on EHR, it is essential that what is set to be one the most important base training datasets for algorithmic development captures details in a manner that makes it work for all, regardless of gender or sexual identity. In the absence of this, AI/machine learning developed using EHR data will remain biased[30]. At the very least, these actions should be considered:

- Creating provisions for allowing patients to voluntarily disclose gender and sexual identity data within EHR design is pivotal. It is important that this information is captured within the demographic section of the record that is more easily and promptly reviewed as opposed to a note added elsewhere that might be missed when switching providers[31]. Creating separate and precise categories to capture these identities (separately for biological sex and gender identity) as proposed by the 2013 recommendations of the World Professional Association for Transgender Health[32] is essential.

- Creating precise definitions and terminologies for different categories of gender and sexual identities and socialising their understanding within the human elements of the system is also important. Such glossaries already exist, but it is important for these to get coded with a shared and common understanding across the system to not create further confusion. In the US, the lack of standardised coding schemes for these categories made it near impossible to exchange this information across EHRs[33].

- Expanding EHRs to include specific forms for making anatomical inventories, at least for patients identified as non-binary in the demographic questioning, could help capture organ diversity and develop appropriate tools and population health systems that consider each patient's gender identity, sex assigned at birth, and anatomy[34]. This can more accurately drive any individualised auto-population of history and physical exam templates, and decision support systems built using these inventories as the base can lead to more appropriate treatment plans.

Of course, at the outset, these systems need to be designed in a human-centred manner to not impinge on privacy or security-based concerns of these individuals who have a legitimate interest in keeping their identities undisclosed to avoid harassment and harm. The EHR is merely a digital

artefact that exists within the larger power and social relationships that also exist in a physician's office, so efforts at sensitivity training and correct impartial training of medical professionals is imperative. It is key that every individual has an opportunity to share information about their sexual orientation and gender identity in a welcoming environment that will facilitate important conversations with clinicians who are able to be helpful.

Ensure data representativeness

It is widely accepted that ensuring representative training data can mitigate the risk of biased outputs. For gender and sexual minorities, it might mean undertaking affirmative action to integrate self-identification as a critical step before bundling a dataset as complete. Annotation or labelling processes rely on a subjective judgement by labelers and this process of inference could be biased or easily mistaken. This is especially true for computer vision applications requiring gender classifications/labels, where the inclusion of self-identification as a part of the informed consent process could mitigate subjective biases[35].

On the part of researchers, the responsibility to prioritise these minorities for physical examinations or collection of specimens for specific studies requires mention. While traditionally large cohort studies are believed to mitigate risk of selection bias and to more accurately depict public health trends in the real world, it is increasingly known that even large sample studies can be skewed, resulting in limited data representativeness[36]. This is where efforts by public institutions can prove beneficial, such as the All of Us Research Program by the US National Institute of Health, which targets creating repositories of biospecimens that consist of majority participation from groups that have been historically underrepresented in biomedical research[37].

Leveraging regulation as a tool

There is no substitute for positive state action to protect vulnerable citizens against discrimination of any kind, whether caused by human or digital service providers. This includes a stringent requirement for gender/sex representativeness in the regulatory pathway for approvals or certification for medical devices, especially in the design, clinical trial, and clinical validation stages[38]. For India, which is still in the early stages of defining the regulatory pathway for software as a medical device, these considerations are paramount and there is immense opportunity to creatively adopt or go beyond the best practices of more mature regulatory authorities like the US FDA[39], given that these biases could actually be more detrimental to the Indian context. Some suggestions include:

- Mandating each dataset to be accompanied by a datasheet explaining the characteristics of the dataset (for instance, motivation, composition and collection process) within the regulatory approval process[40]. This can promote transparency and enable stakeholders using these datasets to train their algorithms to make informed choices on bias.

- Creating institutional mechanisms, either by bolstering capacity within regulatory agencies or by relying on independent third parties for algorithmic impact assessments and audits could also help better define the unintended consequences of automation on certain communities. This should become an integral step to the public procurement process for such technologies.

- Developing an iterative approval process whereby algorithms that perform poorly for specific demographic segments can rework the learning process after collecting additional data from these segments to allow flexibility for a fledgling health-tech industry, such as India's, while not compromising on notions of fairness[41]. For digital biomarkers, this might mean guidance to developers to use cross-validation (or out-of-sample testing) techniques and publish those results in the initial application to the regulator[42]. At the regulator level, this can be followed by both generalisation assessments and tests across specific target populations (implying links back to the positive responsibility on the government research institutions to create repositories of under-represented/vulnerable populations that can be used for this purpose among others). This helps establish boundary conditions for that specific biomarker.

While tactical recommendations can get us only so far and prevent immediate harm, for digital health to be truly transformative to redress the existing inequalities and inaccessibilities of the health system, an acute concern for intersectional gender analyses in digital health research is imperative. This will help tailor research and commensurate products effectively towards specific subgroups, especially in terms of user experience and the optimisation of desired outcomes. Policy regulations, public investments, and evolving ethical considerations governing all stages of the algorithmic value chain, from data generation to AI application, are tools that can propel this goal. This, however, is vastly dependent on a systems approach to viewing the problem, and more importantly, an ecosystem approach to engaging and leveraging all relevant tools in ensuring a safe and efficient healthcare experience for all vulnerable populations in an increasingly digitised world.

Endnotes

1. India Health, Digital Healthcare in India, 2020, New Delhi, Informa Markets, https://www.indiahealth-exhibition.com/content/dam/Informa/indiahealth-exhibition/en/downloads/Digital health report 2020.pdf.

2. Larissa Jennings et al., "Influence of Mhealth Interventions on Gender Relations in Developing Countries: A Systematic Literature Review", *International Journal for Equity in Health* 12, no. 1 (2013), doi:10.1186/1475-9276-12-85.

3. Vijay Kumar Chattu et al., "Fulfilling the Promise of Digital Health Interventions (DHI) to Promote Women's Sexual, Reproductive and Mental Health in the Aftermath of COVID-19", *Reproductive Health* 18, no. 1 (2021), doi:10.1186/s12978-021-01168-x.

4. Tiffany K. Roberts et al., "Barriers to Quality Health Care for the Transgender Population", *Clinical Biochemistry* 47, no. 10-11 (2014): 983-87, doi: 10.1016/j.clinbiochem.2014.02.009.

5. Shamayeta Bhattacharya et al., "Studying Physical and Mental Health Status among Hijra, Kothi and Transgender Community in Kolkata, India", *Social Science & Medicine* 265 (2020): 113412, doi: 10.1016/j.socscimed.2020.113412.

6. Clemens Scott Kruse et al., "The Use of Electronic Health Records to Support Population Health: A Systematic Review of the Literature", *Journal of Medical Systems* 42, no. 11 (2018), doi:10.1007/s10916-018-1075-6.

7. T. A. Holt, et al., "Identification of Undiagnosed Diabetes and Quality of Diabetes Care in the United States: Cross-sectional Study of 11.5 Million Primary Care Electronic Records", *CMAJ Open* 2, no. 4 (2014), doi:10.9778/cmajo.20130095.

8. Truyen Tran et al., "Risk Stratification Using Data from Electronic Medical Records Better Predicts Suicide Risks than Clinician Assessments", *BMC Psychiatry* 14, no. 1 (2014), doi:10.1186/1471-244x-14-76.

9. Evidence also shows that even when sex and gender identity are recorded in the EHR, they are stored in a variety of ways. Some providers or sites would record it within demographic details while others in social history, or some even leave it as an extra note somewhere in the chart.

10. Edward J. Callahan et al., "Introducing Sexual Orientation and Gender Identity Into the Electronic Health Record", *Academic Medicine* 90, no. 2 (2015): 154-60, doi:10.1097/acm.0000000000000467.

11. Department of Information Technology, Ministry of Communications and Information Technology, Government of India, http://egovstandards.gov.in/sites/default/files/MDDS_Demographic_Ver1.1.pdf.

12. National LGBT Health Education Centre, *Collecting Sexual Orientation and Gender Identity Data in Electronic Health Records*, The Fenway Institute, 2016, https://www.lgbtqiahealtheducation.org/wp-content/uploads/Collecting-Sexual-Orientation-and-Gender-Identity-Data-in-EHRs-2016.pdf.

13. Chris Grasso et al., "Optimizing Gender-affirming Medical Care through Anatomical Inventories, Clinical Decision Support, and Population Health Management in Electronic Health Record Systems", *Journal of the American Medical Informatics Association*, 2021, doi:10.1093/jamia/ocab080.

14. Chris Grasso et al., "Using Sexual Orientation and Gender Identity Data in Electronic Health Records to Assess for Disparities in Preventive Health Screening Services", *International Journal of Medical Informatics* 142 (2020): 104245, doi:10.1016/j.ijmedinf.2020.104245.

15. Lester Darryl Geneviève et al., "Structural Racism in Precision Medicine: Leaving No One behind", *BMC Medical Ethics* 21, no. 1 (2020), doi:10.1186/s12910-020-0457-8.

16. Alyson J. Mcgregor et al., "How to Study the Impact of Sex and Gender in Medical Research: A Review of Resources", *Biology of Sex Differences* 7, no. S1 (2016), doi:10.1186/s13293-016-0099-1.

17. Lisa A. Martin et al.,. "The Experience of Symptoms of Depression in Men vs Women", *JAMA Psychiatry* 70, no. 10 (2013): 1100, doi:10.1001/jamapsychiatry.2013.1985.

18. Lmar M. Babrak et al., "Traditional and Digital Biomarkers: Two Worlds Apart?", *Digital Biomarkers* 3, no. 2 (2019): 92-102, doi:10.1159/000502000.

19. Christopher W. Snyder et al., "The Best Digital Biomarkers Papers of 2017", *Digital Biomarkers* 2, no. 2 (2018): 64-73, doi:10.1159/000489224.

20. J. Torous et al., "New Dimensions and New Tools to Realize the Potential of RDoC: Digital Phenotyping via Smartphones and Connected Devices", *Translational Psychiatry* 7, no. 3 (2017), doi:10.1038/tp.2017.25.

21. "Indian Women Less Likely than Men to Own a Mobile, but Are Catching up", *Business Standard*, February 7, 2021, https://www.business-standard.com/article/politics/indian-women-less-likely-than-men-to-own-a-mobile-but-are-catching-up-121020700857_1.html.

22. Rachel Metz, "AI Software Defines People as Male or Female. That's a Problem", *CNN*, November 21, 2019, https://edition.cnn.com/2019/11/21/tech/ai-gender-recognition-problem/index.html.

23. As one blogpost on Medium's Towards Data Science describes it: Taking the example of Berlin Transportation Company's (BVG) 21% ticket discount to 'real' women on Equal Pay Day, where the BVG staff could make decisions about the gender identity of its commuters, it could be reasonably assumed that automation using a camera and AI could very well erase the trans-identity of the commuters by classifying them one way or another; Zachary Hay, "Towards Trans-Inclusive AI," *Towards Data Science*, May 13, 2019, https://towardsdatascience.com/towards-trans-inclusive-ai-a4abe9ad4e62.

24. Os Keyes, "The Misgendering Machines", *Proceedings of the ACM on Human-Computer Interaction* 2, no. CSCW (2018): 1-22, doi:10.1145/3274357.

25. Jaden Urbi, "Some Transgender Drivers Are Being Kicked off Uber's App", *CNBC*, August 8, 2018, https://www.cnbc.com/2018/08/08/transgender-uber-driver-suspended-tech-oversight-facial-recognition.html.

26. Musaed Alhussein et al., "Automatic Gender Detection Based on Characteristics of Vocal Folds for Mobile Healthcare System." *Mobile Information Systems* (2016): 1-12, doi:10.1155/2016/7805217.

27. For instance, a recent 2021 study proposed a model for automatic gender detection/classification based on twitter profiles intended to serve as an easy pipeline to provide demographic prediction to other social media-based health cohort studies because demographic details like sex/gender is not explicitly known for social media users. However, the same study noted in its limitations that, "as it is estimated that less than 0.5% of the US population are considered as transgender their contributions to the classification performance should not be significant in this work"; Yuan-Chi Yang et al., "Automatic Gender Detection in Twitter Profiles for Health-related Cohort Studies" (2021), doi:10.1101/2021.01.06.21249350.

28. Vivek Divan et al., "Transgender Social Inclusion and Equality: A Pivotal Path to Development", *Journal of the International AIDS Society* 19 (2016): 20803, doi:10.7448/ias.19.3.20803.

29. Liza Khan, "Transgender Health at the Crossroads: Legal Norms, Insurance Markets, and the Threat of Healthcare Reform", *Yale Journal of Health Policy and Ethics* 11 (2011).

30. Milena A. Gianfrancesco, et al., "Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data", *JAMA Internal Medicine* 178, no. 11 (2018): 1544, doi:10.1001/jamainternmed.2018.3763.

31. Rajiv Leventhal, "The Many Layers of Healthcare's EHR Gender Identity Problem." Healthcare Innovation, May 14, 2018. https://www.hcinnovationgroup.com/clinical-it/article/13029566/the-many-layers-of-healthcares-ehr-gender-identity-problem.

32. Deutsch MB et al., "Updated recommendations from the world professional association for transgender health standards of care", *American Family Physician* 87, no. 2 (2013): 89-93, PMID: 23317072.

33. Madeline B. Deutsch et al., „Electronic Health Records and Transgender Patients—Practical Recommendations for the Collection of Gender Identity Data", *Journal of General Internal Medicine* 30, no. 6 (2015): 843-47, doi:10.1007/s11606-014-3148-7.

34. An anatomical inventory is a record of a patient's medical transition history and current anatomy and helps record/update into the chart the organs each individual patient has at any given point in time; Grasso et. al. "Optimizing Gender-affirming Medical Care through Anatomical Inventories, Clinical Decision Support, and Population Health Management in Electronic Health Record Systems."

35. Morgan Klaus Scheuerman et al., „How Computers See Gender:, *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (2019): 1-33, doi:10.1145/3359246.

36. Katherine M. Keyes et al, "UK Biobank, Big Data, and the Consequences of Non-representativeness", *The Lancet* 393, no. 10178 (2019): 1297, doi:10.1016/s0140-6736(18)33067-8.

37. "The "All of Us" Research Program", *New England Journal of Medicine* 381, no. 7 (2019): 668-76, doi:10.1056/nejmsr1809937.

38. Agostina J. Larrazabal et al., "Gender Imbalance in Medical Imaging Datasets Produces Biased Classifiers for Computer-aided Diagnosis", *Proceedings of the National Academy of Sciences* 117, no. 23 (2020): 12592-2594, doi:10.1073/pnas.1919012117.

39. Office of the Commissioner, U.S. Food and Drug Administration, Government of the United States of America, https://www.fda.gov/science-research/womens-health-research/understanding-sex-differences-fda.

40. Timnit Gebru et al., "Datasheets for Datasets" [Working Paper] *ArXiv* (2020), https://arxiv.org/abs/1803.09010.

41. Irene Y. Chen et al., "Why Is My Classifier Discriminatory?", *Proceedings of 32nd Conference on Neural Information Processing Systems* (2018).

42. Karen D. Davis et al., "Discovery and Validation of Biomarkers to Aid the Development of Safe and Effective Pain Therapeutics: Challenges and Opportunities", *Nature Reviews Neurology* 16, no. 7 (2020): 381-400, doi:10.1038/s41582-020-0362-2.

Challenging the Dogmatic Inevitability of Extraterritorial State Surveillance

Arindrajit Basu

Networked data trails define the lives of individuals and communities in many ways, and consequently also serve as ripe prospects for state surveillance. Digital surveillance is now encoded into the governance architectures across political and economic systems.

Largely, surveillance and its legal restraints have revolved around domestic surveillance exerted by states on their citizens. This includes the all-encompassing domestic mass surveillance systems operationalised in authoritarian regimes like China that contain few safeguards constraining their practices in legislation or constitutional architecture[1], as well as limited surveillance techniques increasingly being deployed by several democracies to target suspected criminals, and allegedly, individuals and communities dissenting against the state[2]. Several democracies, including India, boast robust constitutional protection of fundamental rights but are yet to undertake a thorough reform of their legislation enabling domestic surveillance that predate the digital era[3]. The scrutiny of domestic surveillance practices is important, but at the same time, extraterritorial surveillance by countries, including democracies, continues largely unchecked. Several countries engaging in widespread extraterritorial surveillance, including the US and UK, boast robust legislative protections for their citizens against state surveillance but impose woefully few restraints when their intelligence agencies conduct surveillance abroad.

From a global standpoint, both domestic and foreign surveillance warrant attention, scrutiny, and critique. The technology and ideology of the Chinese surveillance state is being exported to other countries to some degree, warranting possible cause for concern[4]. However, recent scholarship has questioned this notion, arguing that "China and other autocracies cannot simply will into existence oversees replicas of their surveillance states[5]." Empirical work found limited evidence that China was a leading driver of digital repression in other countries[6].

Some democracies that engage in extraterritorial surveillance, like the US, possess far greater surveillance capabilities than the rest of the world. That they choose to restrain these capabilities when deploying them against their citizens but not when acting abroad should be a matter of concern for the international community. As the US plans a 'summit for democracy'[7]to be held later this year, the democratic legitimacy of foreign surveillance practices must be subjected to scrutiny. As of now, there is an acceptance of inevitability around extraterritorial surveillance practices by policymakers and the judiciary around the world, boxed in with the vast array of arbitrary measures shielded by dogma around national security concerns.

Law and State Practice on Extraterritorial Surveillance

Surveillance is the focused systematic and routine gathering of information through the identification, tracking, monitoring and analysis of data belonging to individuals, organisations, or communities[8]. In the cyber realm, surveillance can either be targeted or untargeted. Targeted surveillance is the covert collection of metadata, communications, and conversations after

suspicion against a specific target has been established. When states do not have the capabilities to conduct targeted surveillance themselves or seek to avoid political or diplomatic costs, they rely on private commercial players who provide sophisticated spyware to government clients to infiltrate a target's device[9]. Firms like Israel based NSO Group (that allegedly markets the Pegasus spyware), Germany based Finfisher Group, and the Italy based Hacking Team are the largest players in this sector[10].

States also engage in untargeted surveillance. Commonly known as 'mass surveillance', this involves the aggregation of vast amounts of data collected through the systematic 'bulk' monitoring of online actions and behaviour by individuals and communities online[11]. While most states have clouded their surveillance practices in a veil of opacity, revelations by Edward Snowden in 2013 provide some unique insight into the programmes run by the US National Security Agency (NSA). The NSA ran several programmes that collect information containing both metadata and content. Extraterritorially, through its relationship with telecom firm AT&T, the NSA was able to access data through international fibre optic cables, thereby targeting several countries, including France, Germany, Greece, Italy, Japan, Mexico, South Korea and Venezuela[12]. Given the central role played by the US in internet architecture, the NSA was also able to access data travelling through choke points on US territory[13]. In addition, to 'upstream collection,' which relies on access to fibre optic cables and other infrastructure, the NSA was able access the content of communications from the servers of the nine largest internet companies in the world through its infamous PRISM programme[14].
After bulk collection, the data was processed through a range of data mining techniques that help identify potential suspects and conduct targeted surveillance. Through a process known as data-chaining, recorded events were mapped onto a set of topographical patterns and subsequently filtered by algorithmic processes to highlight suspicious patterns through analytical programmes like XKeyScore[15].

Spurred by backlash to the Snowden revelations, some countries have developed legal regimes requiring state agencies to abide by the principles of legality, proportionality and necessity when conducting domestic surveillance[16], ensuring oversight by independent bodies and granting legal remedies to individuals unfairly targeted[17]. These legal regimes stem from constitutional guarantees that protect individual privacy, such as the Fourth Amendment in the US[18].However, most of these safeguards do not apply to extraterritorial surveillance and information gathering by intelligence agencies has been an arena where the "principles of a democratic state have been applied less conscientiously[19]." For example, Section 702 of the 2008 Foreign Intelligence Surveillance Amendment Acts affords differing levels of protection to US persons and non-US persons from the NSA's surveillance programmes[20]. US persons, including American citizens and non-citizens permanently residing in the country, can be targeted only if there is both probable cause to believe that the individual is a foreign agent and a warrant is issued by the Foreign Intelligence Surveillance Court (FISC). Non-US persons, on the other hand, can be the target of surveillance under a much lower 'reasonable belief' standard that does not require probable cause or a warrant from the FISC.

Still, targeted interception of foreign communications is at least regulated by specific laws while mass surveillance remains largely unrestrained. In 2015, 27 of the then 28 EU member states had laws for targeted interception, but only five (France, Germany, UK, the Netherlands,

and Sweden) had specific legislation regulating and restraining the bulk interception of foreign communications[21]. In the US, the Reagan-era Executive Order 12333 continues to empower the president to conduct surveillance activities abroad[22]. In response to the public pressure exerted as a result of the Snowden revelations, former President Barack Obama issued the Presidential Policy Directive-28 (PPD-28)[23]. The PPD-28 sought to bring transparency and legitimacy to the signals intelligence process by recognising the dignity and legitimate privacy interests of individuals regardless of nationality or residence[24], and introduce safeguards drawn from data protection frameworks, including data minimisation, limits on use, and retention. On the other hand, it endorsed bulk collection as necessary in a networked world and included a broad range of instances, including espionage, terrorist threats to the US, threats due to weapons proliferation, threats to US or allied armed forces, and international criminal threats[25]. While these specified purposes are exhaustive, 'terrorist threats' and 'cybersecurity' are broad enough to enable a variety of purposes for mass surveillance with no opportunity for legal scrutiny by the individuals and communities at the receiving end of these bulk collection practices[26].

The UK's extraterritorial surveillance programme, laid out through the Regulation of Investigatory Powers Act, 2000 (now Investigatory Powers Act, 2016), also draws a distinction between "internal" and "external" communications[27]. Domestic surveillance needs individualised warrants, whereas overseas data merely requires a bulk targeting warrant with no cap on the quantity of data that may be collected. Restraints on this largely echo PPD-28-strong procedural safeguards are emphasised but the notion that bulk interception is essential for preserving national security interests is placed on a pedestal[28].

Regimes that undertake draconian surveillance at home do not restrain themselves when it comes to extraterritorial surveillance if they possess the capabilities to do so. China, for example, has allegedly used mobile phone networks in the Caribbean to conduct surveillance on US mobile phone subscribers[29]. However, sans a Snowden-like whistleblower emerging from China and revealing the full extent of extraterritorial surveillance, our knowledge and understanding of its strategy remains piecemeal and limited. For now, international law must question the practices that are widely known, not least because of the inherent paradox between domestic protection and extraterritorial overreach.

Whither International Law?

Yet, international law has remained strikingly quiet on this subject and failed to cast binding obligations or influence state practice through treaty law or customary international law[30]. The reason for this is the unwavering belief that the ability to gather intelligence is essential to statecraft and national security[31]. According to this notion, consenting to international obligations that restrain specific activities in this domain will end up harming the state's core interests.

In the absence of hard law, a body of soft law interpreting existing treaty regimes to constrain surveillance has emerged. Article 17 of the International Covenant on Civil and Political Rights (ICCPR) protects the right to privacy against 'unlawful' and arbitrary interference[32]. The UN Human Rights Committee has stated that surveillance can only take place on the basis of a law that is well-defined, accessible to the public, and pursues legitimate aims[33]. This has been affirmed by the UN General Assembly in resolutions 68/167 and 69/166, which urge states to

ensure transparency and accountability of surveillance practices[34]. While the resolutions signify some political commitment to an issue, scholars have described them as products of a 'fake consensus'[35] —a facade of commitment to the rules-based order maintained by states rather than an undertaking to genuinely improve state behaviour. The failure of states to satisfactorily amend their surveillance practices to comply with international human rights law has been recognised by Michelle Bachelet, the UN High Commissioner for Human Rights[36], and David Kaye, the former UN Special Rapporteur on Freedom of Expression and Opinion[37].

National, supranational and constitutional courts have taken up the question of bulk surveillance[38]. However, most of these judgements have revolved around the question of domestic surveillance or scenarios where the rights of their citizens were at stake. Even the EU's progressive courts—Court of Justice of the European Union (CJEU) and the European Court of Human Rights (ECtHR)—have been quiet on the question of extraterritorial surveillance. Only three cases from the ECtHR referred to the conduct of foreign surveillance by states[39], but none of the judgements issued a clear and unambiguous interpretation of whether and how extraterritorial surveillance of non-citizens will violate provisions of the European Convention of Human Rights[40]. More problematically, they endorse the notion that the interception of foreign communications falls within a state's 'margin of appreciation' and that legislative protection that discriminates based on the target's geographical location is valid[41]. A lack of clarity on the legality of extraterritorial surveillance only strengthens the isolated US (and Israeli) position that human rights obligations do not apply to non-citizens located outside their territory, jurisdiction or 'effective control'[42]. Even though this position is based on a dubious reading of Article 2(1) of the ICCPR and has been rebutted by scholars [43]and successive reports of the UN Human Rights Committee[44], the justification continues to legitimise US surveillance practices.

The European Court of Human Rights could do a lot more to drive norms delegitimising and constraining extraterritorial surveillance and could derive some lessons from a 2020 decision of the German Constitutional Court[45]. This ruling held that bulk surveillance of non-citizens outside Germany violated the German Basic Law, arguing that protection of this nature reflected Germany's participation in the international community. Essentially, it articulated that the architecture of fundamental rights in German law casts obligation on the German state whenever it exercises 'effective control 'over the rights of individuals regardless of location or nationality[46]. This position on extraterritorial applicability of human rights has also found mention in General Comment No.36 of the UN Human Rights Committee[47] and ECtHR jurisprudence, although the German case marks the only instance where this has been applied to extraterritorial surveillance.

Jurisprudence of both the ECtHR and CJEU suffer from another key flaw. There is an overreliance on procedural safeguards restraining extraterritorial bulk surveillance without going into the costs and benefits of mass surveillance. Procedural safeguards for any practice mean little when its implementation is secretive and cannot be challenged by the intended targets of that very practice. This 'procedural fetishism' [48] is backed by a dogmatic acceptance of the national security necessity of mass surveillance and limited engagement with the social and political costs of these practices. But there is a need to break from this unquestioned consensus.

Questioning National Security and Highlighting Costs

To justify the NSA's bulk targeting program, the agency's former director Keith Alexander told the US House Intelligence Committee that surveillance had helped thwart over 50 incidents since the September 2001 attacks [49]. While Alexander provided some examples of these potential attacks, the exact role played by specific surveillance programmes in preventing an attack was not disclosed. This is a feature of surveillance conversations world over—states use the "trump card"[50] of national security to justify their actions and to keep the evidence backing this claim classified, thereby preventing an open, frank, and transparent public discussion on the issue. In fact, Alexander later admitted that the 50-plot figure was incorrect, and instead only one activity involving the transfer of a small amount of funds to a banned terrorist group in Somalia was foiled[51]. Intelligence officials do not evaluate the effectiveness of specific surveillance programmes as they believe that intelligence work is a "puzzle", and that one piece of this puzzle cannot be measured without evaluating the complete picture[52]. Research by New America Foundation in 2014 argues that traditional targeted intelligence methods, such as tips from local communities, played a far more critical role than foreign surveillance, which played a role in only 4.4 percent of the terrorist cases examined[53]. These findings and other research making the same case [54] have done little to alter the status quo and the unquestioned dogma around national security-related benefits of mass surveillance, which continue to drive thinking among governments and courts alike. This dogmatic approach also betrays a reductive understanding of security as the mere absence of harm. A more holistic bottom-up approach will also assess whether individuals and communities feel secure in the integrity of their online communications[55]. This approach will question whether the absence of terrorist attacks or the weeding out of terrorist threats meets the threshold of a secure life if it leads to individuals and communities choosing to use online communication channels less frequently for fear of being unwittingly pried upon by foreign governments.

The systematic padding of national security benefits in state-driven discourse is accompanied by a consistent erasure of costs, to human dignity and political life. The threat to human dignity stems from datafication—an individual's data is manipulated and formatted to extract a pattern about that person's world, such that they no longer speak for themselves[56]. Instead, they are massaged into various categories that the subject does not understand or exercise control over[57]. This results in gross power asymmetry between those controlling and using surveillance technology, and those at the receiving end of it. Mass surveillance is akin to an analogue situation where an unknown entity many continents away could collect the personal diaries of all individuals and store it in a safe, always retaining the power to unlock the safe and read the content of the diary whenever needed through targeted surveillance. The individual or community is powerless due to gaps in knowledge regarding the extent of scrutiny on their daily existence combined with a lack of political or legal remedy. Algorithmic processing of data gathered through mass surveillance is equally harmful. Discrimination is an inherent possibility of the cultural bias of the developer generating the training data and source code for the algorithm getting quantified through hidden layers in a manner that confirms certain practices as belonging to the realm of the 'other'[58]. Certain lifestyles and cultural aspirations are thereby branded 'suspicious' for reasons that are 'black boxed' and subsequently targeted by algorithmically-driven analytical programmes like XKeyscore, thus denying them respectability and agency[59]. Individuals then change their behaviour to avoid being branded as suspicious and to refrain from speech that might cross the acceptable limits, as defined by the watchers. Thus, surveillance has political costs for civic life.

An oft-repeated argument justifying discrimination based on geography is that states cannot quash dissent or exact police powers outside their territory, thus negating the political costs of surveillance[60]. This argument is flawed for several reasons. First, the architecture of control in the case of global surveillance stems from uncertainty-individuals do not know whether, why and how their intimate daily activities are being monitored and how this information may be used. It is incorrect to say that states lack police powers abroad; lethal drone or air strikes are often used to take out suspected terrorists. While these actions too are governed by strict laws and procedures, they may further compound anxiety and uncertainty, as individuals seek to avoid being cast into labels that might cause them to become targets.

Second, in the case of domestic surveillance, citizens have the right to sue the state in court or vote governments out of power. Individuals abroad are helpless in this regard, forced into an acceptance of the global surveillance state and life with it.

Third, this underestimates the pervasive nature of intelligence sharing among countries. The Snowden revelations showed that the NSA skirted domestic British law by sharing information on British citizens with their counterparts in the UK. Although Snowden uncloaked the intelligence-sharing agreements between the 'Five Eyes' (Australia, Canada, New Zealand, the UK and the US), such deals are often confidential and not publicly scrutinised, being brokered directly between the relevant ministries or agencies. Therefore, data collection by a foreign government could end up in the hands of the individual's own government, without the individual's knowledge. Relatedly, surveillance strategies and methods devised and bred in one part of the world can easily be transferred elsewhere. In September 2021, news reports suggested that former NSA intelligence analysts started working for the United Arab Emirates monarchy to engage in surveillance of militants, domestic dissenters, and foreign governments and citizens[61]. Crucially, these surveillance practices also targeted US citizens, ironically using methods that were incubated by the US for extraterritorial deployment.

Finally, there is an asymmetry of capabilities between states. The status quo legitimises a scenario where, by default, the citizens of countries with lesser capabilities will face greater intrusions in their privacy than those of more powerful states. This reality violates sovereign equality, which is a cardinal tenet of international relations, and undermines the trust that governments and citizens repose in each other[62].

Towards a Norm Against Surveillance

Modern foreign surveillance need not become an inevitability if the international community takes steps to act against it and operationalise legal and normative frameworks condemning it. One might question the extent to which global normative frameworks could change state practices, but history is witness to its value. Until the signing of the Kellogg-Briand Pact in 1928, the act of war was legal and an acceptable method of resolving disputes between states[63]. As with foreign surveillance today, states perceived war as an essential tool of statecraft and an instrument of the rule of law. The outlawing of war through the 1928 pact and the subsequent prohibition on the threat or use of force in Article 2(4) of the UN Charter brought about a radical transformation in how states perceived and conducted their affairs with other states[64]. Conquest for the purpose of acquiring resources was common but is now a rare occurrence. There were

also other factors at play, such as the destructive capabilities of modern weapons and increasing economic interdependence[65]. Naturally, the causal connection between these factors and the use of force today should not be ignored. The occlusion of the threat of war from the table of state relations allowed alternate mechanisms of cooperation and contestation—such as economic sanctions, confidence-building measures, and development partnerships—to grow and benefit global governance. Norms by themselves do not dramatically alter state behaviour but they provide a framework for enforcing incentives that do so[66].

A similar norm outlawing the use of surveillance as an unquestioned tool of statecraft must be operationalised. This does not mean that foreign surveillance will always be illegal. The use of force is justified by exceptions incorporated into the UN Charter—self-defence and UN Security Council authorisation. There could be similar exceptions that states consent to, justifying the use of foreign surveillance. Foreign surveillance should not be accepted by default. Altering the normative landscape on foreign surveillance will change state incentives in several ways. In the short run, it will lend weight to economic sanctions, such as the call by UN Special Rapporteurs for a moratorium on the sale of export control technologies[67]. States whose citizens are at the receiving end of foreign surveillance will be able to ask questions of, and demand accountability from, countries conducting these practices. Moving forward, civil society voices around the world will be empowered to challenge practices by their governments and others.

Given the existing entrenched dogma around surveillance, a key challenge is operationalising the norm in the first place. States remain reluctant to take any concrete steps in this direction. Again, history shows that key rules of international law were initiated and promoted by several non-state actors, or 'norm entrepreneurs'. The diplomatic process outlawing war was conceptualised and initiated by a Chicago-based commercial lawyer who worked with several burgeoning peace movements at the time[68]. Other examples include private actors like DuPont, which played a major role in the Montreal Protocol; non-governmental organisations like the International Committee of the Red Cross that galvanised and helped operationalise the Geneva Conventions; and inter-governmental organisations like the Asian-African Legal Consultative Organisation that negotiated several key provisions in the UN Convention on the Law of the Sea[69].

Challenging foreign surveillance by states needs an ecosystem of actors to come together, transcending the artificial segregation between self-contained groups of 'rights-holders' and 'rights-violators'[70]. The politics of surveillance is not a binary tussle between the 'intelligence agencies' and the 'rest of the world', operating as two homogenous entities[71]. The coders working for the NSA or designing the algorithms have multiple identities and relationships. Their families and friends, and they themselves, are subjects of the surveillance process. A norm against surveillance can be operationalised once the costs are brought to the fore for shared lived experiences, rather than opaque security metrics, galvanises meaningful public discourse. Several multistakeholder groupings working on cybersecurity and digital rights such as the Freedom Online Coalition, Global Network Initiative, and the Global Forum on Cyber Expertise bring together diverse perspectives from governments, civil society and the private sector, and could act as the catalyst for driving such a norm through the institutional international law framework.

The project of international law and the rules-based global order was conceived with lofty ideals, that the strong restrain themselves to ensure that the weak need not endure what they must.

The preservation of the cardinal tenets of international law rests most on the shoulders of those who seek to promote it. Rather than wax eloquent about the harmful practices of other actors, it is imperative that countries look beyond dogmatic assumptions to restrain their own practices and catalyse global norms that protect the rights of individuals and communities, regardless of nationality or geography.

Endnotes

1. Ross Andersen," The panopticon is already here," *The Atlantic*, September 2020. https://www.theatlantic.com/magazine/archive/2020/09/china-ai-surveillance/614197/.

2. Steve Feldstein, *The rise of digital repression: How technology is reshaping power, politics, and resistance* (New York: Oxford University Press,2021).

3. Tanmay Singh and Anushka Jain," Surveillance reform is the need of the hour," *The Hindu*, July 10,2021, https://www.thehindu.com/opinion/op-ed/surveillance-reform-is-the-need-of-the-hour/article35414371.ece.

4. Steven Feldstein," The global expansion of AI surveillance," *Carnegie Endowment for International Peace*, September 17,2019, https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847.

5. Christopher Walker, Shanthi Kalathil, and Jessica Ludwig, "The cutting edge of sharp power," *Journal of Democracy* 31, no 1 (2020):130-31.

6. Feldstein, *The rise of digital repression*, pp.50.

7. Sriram Lakshman, " Biden to host democracy summit in December," *The Hindu*, August 11,2021, https://www.thehindu.com/news/international/biden-to-host-democracy-summit-in-december-virtually/article35857080.ece.

8. Torin Monahan, Surveillance in the time of insecurity (New Brunswick: Rutgers University Press,2010), pp.8. David Lyon, *Surveillance Studies* (Polity Press,2007), pp.14.

9. The Citizen Lab," Communities@Risk: Targeted digital threats against digital society," *The Citizen Lab*, November 11,2014, https://targetedthreats.net/media/1-ExecutiveSummary.pdf.

10. The Citizen Lab, "Communities@Risk: Targeted digital threats against digital society".

11. Roger A. Clarke, 'Information Technology and Dataveillance' Comm. Acm 31 (1998): 498; Margaret Hu, 'Small Data Surveillance v. Big Data Cybersurveillance' (2015) *Pepp.Law Rev* 42:776-777.

12. Glenn Greenwald, *No Place to Hide: Edward Snowden, the NSA and the Surveillance State*, (New York: Metropolitan Books, 2015).

13. "NSA slides explain the PRISM data-collection program," *The Washington Post*, June 6,2013, https://www.washingtonpost.com/wp-srv/special/politics/prism-collection-documents/.

14. Glenn Greenwald and Evan Macaskill, "NSA Prism program taps into user data of Apple,Google and others," *The Guardian*, June 7,2013, https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data.

15. George Lucas, *Ethics and Cyber Warfare: The quest for responsible security in the age of digital warfare* (New York: Oxford University Press, 2016), pp.144.

16. Marcin Rojszczak," Extraterritorial bulk surveillance after the German BND Act Judgment," *European Constitutional Law Review 17*(1) (2021):53.

17. Commissioner for Human Rights, *Democratic and effective oversight of national security services*. Brussels, 2015.

18. Laura Donohue, 'The Fourth Amendment in a Digital World', *NYU Annual Survey of American Law 71* (2017): 533.

19. Rojszczak, "Extra-territorial bulk surveillance after the German BND Act".

20. Eliza Watt, "The right to privacy and the future of mass surveillance," *The International Journal of Human Rights,* Volume 21 No 7 (2017):773.

21. EU Agency for Fundamental Rights, *Surveillance by intelligence agencies: Fundamental rights safeguards and remedies in the EU-Mapping member states legal frameworks*, Brussels*:* 2017.

22. Executive Order No. 12333, 3 C.F.R. 200 (December 4,1981).

23. Office of the Press Secretary, The White House, US Government, "Presidential Policy Directive No.28 Signals Intelligence Activities," January 27, 2014, https://obamawhitehouse.archives.gov/the-press-office/2014/01/17/presidential-policy-directive-signals-intelligence-activities.

24. Presidential Policy Directive No.28 Signals Intelligence Activities.

25. Presidential Policy Directive No.28 Signals Intelligence Activities.

26. Presidential Policy Directive No.28 Signals Intelligence Activities,

27. Investigatory Powers Act, 2016, §136(4)

28. Government of UK, *Government Response to the ISC Report on Privacy and Security*. London:2015. https://isc.independent.gov.uk/wp-content/uploads/2021/01/20151208_Privacy_and_Security_Government_Response.pdf.

29. Stephanie Kirchgaessner, " Revealed: China suspected of spying on Americans via Carribean phone networks," *The Guardian*, Dec 15,2020, https://www.theguardian.com/us-news/2020/dec/15/revealed-china-suspected-of-spying-on-americans-via-caribbean-phone-networks.

30. Craig Forcese, "Spies without borders: International Law and Intelligence Collection," Journal *of National Security Law and Policy* 5 (2016): 205.

31. Buchan, *Cyber espionage and international law*, pp.6.

32. Sarah Joseph and Melissa Castan, *The International Covenant on Civil and Political Rights,3rd ed* (Oxford: Oxford University Press, 2013), pp.533

33. United Nations General Assembly, *Report of the Office of the United Nations High Commissioner for Human Rights the Right to Privacy in the Digital Age*, Geneva,Human Rights Council,2014.

34. Eliza Watt,*State sponsored cyber surveillance: The right to privacy of communications and international law* (Cheltenham: Edward Elgar Publishing,2020), pp.130.

35. Stephen Schwebel" The effect of resolutions of the UN General Assembly on customary international law" American Society of International Law" (1979):301 in Watt,*State Sponsored Cyber Surveillance*,pp.131.

36. United Nations General Assembly, *Report of the Office of the United Nations High Commissioner for Human Rights the Right to Privacy in the Digital Age*, Geneva:Human Rights Council,2018.

37. United Nations General Assembly, *Surveillance and human rights: Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression*, Geneva: Human Rights Council,2019.

38. Watt, *State sponsored cyber surveillance.*

39. *Weber and Saravia v Germany* App no 54934/00 (June 29,2006), *Centrum for Rattvisa v Sweden* App no 35252/08 (19 June 2018), *Big Brother Watch and Others v the United Kingdom* App no 58170/13.

40. David Cole and Federico Fabbrini, "Bridging the transatlantic divide? The United States, the European Union, and the protection of privacy across borders," *International Journal of Constitutional Law* 14(2016): 220; Rojsczcak, "Extra-territorial bulk surveillance after the German BND Act".

41. Watt, State sponsorer cyber surveillance, pp.142.

42. Marko Milanovic," Harold Koh's legal opinions on the US position on the extra-territorial application of human rights treaties," *EJIL Talk*, March 7,2014, https://www.ejiltalk.org/harold-kohs-legal-opinions-on-the-us-position-on-the-extraterritorial-application-of-human-rights-treaties/.

43. Peter Margulies, Surveillance By Algorithm: The NSA, Computerized Intelligence Collection, and Human Rights, *Florida Law Review*, Volume 68 No 4 (2016):1055.

44. UNGA, *Report of the Office of the United Nations High Commissioner for Human Rights the Right to Privacy in the Digital Age*, UN Doc. A/HRC/27/37,2014.

45. BVerfG 19 May 2020, 1 BvR 2835/17

46. Basak Cali, " Has 'control over rights doctrine' for extra-territorial jurisdiction come of age? Karlsruhe too,has spoken,now it's Strasbourg's turn," *EJIL Talk*,July 21,2020, https://www.ejiltalk.org/has-control-over-rights-doctrine-for-extra-territorial-jurisdiction-come-of-age-karlsruhe-too-has-spoken-now-its-strasbourgs-turn/.

47. United Nations Human Rights Committee, *General Comment No.36 Article 6 *Right to Life*,* Geneva: Human Rights Committee,2018.

48. Monika Zalnierute, " A dangerous convergence: The inevitability of mass surveillance in European Jurisprudence," *EJIL Talk,* June 4,2021, https://www.ejiltalk.org/a-dangerous-convergence-the-inevitability-of-mass-surveillance-in-european-jurisprudence/.

49. Bailey Cahall, Peter Bergen and David Sterman,"Do NSA's bulk surveillance programs stop terrorists," *New America*, January 13,2014,https://www.newamerica.org/international-security/policy-papers/do-nsas-bulk-surveillance-programs-stop-terrorists/.

50. Jennifer A. Chandler, "Personal Privacy versus National Security: Clarifying and Reframing the trade-off" in Ian Kerr, Valerie Stevens and Carole Lucock (eds) *On the identity trail: Anonymity, Privacy and Identity in a Networked Society* (Oxford: Oxford University Press, 2009) pp.121-138.

51. Keith Alexander, Evidence before the United States Senate Committee on the Judiciary. In: Continued Oversight of the Foreign Intelligence Surveillance Act, Washington DC: 2013.

52. Michelle Cayford & Wolter Pieters, " The ineffectiveness of surveillance technology: What intelligence officials are saying," *The Information Society,* 34(2), 2018, p 88.

53. Peter Bergen et al, "Do NSA's bulk surveillance programs stop terrorists," *New America Foundation*, 2014, https://www.jstor.org/stable/resrep10476?seq=1#metadata_info_tab_contents.

54. Jennifer Stisa Granick, "Mass spying isnt just intrusive-it's ineffective," *Wired*, February 03,2017, https://www.wired.com/2017/03/mass-spying-isnt-just-intrusive-ineffective/.

55. Anja Kovacs and Dixie Hawtin, "Cybersecurity,surveillance and human rights," *Global Partners Digital,* January 31st 2013,https://www.gp-digital.org/publication/second-pub.

56. John Cheney-Lipold,*We are data: Algorithms and the making of our digital selves* (New York: NYU Press,2017),pp.14.

57. Louise Amoore, "Data Derivatives:On the Emergence of a Security Risk Calculus for Our Time," (2011) *Theory, Culture and Society*,28(6) (2011):24.

58. Claudia Aradau and Tobias Blanke, "The (Big) Data-security assemblage: Knowledge and critique," *Big Data & Society*, 2(2) (2015):1.

59. Frank Pasquale, *Black Box Society: The Secret Algorithms That Control Money and Information* (New York: Harvard University Press,2015).

60. Peter Swire, Jesse Woo, Deven Desai," The important, justifiable, and constrained role of nationality in foreign intelligence surveillance," *Lawfare*, January 11,2019, https://www.lawfareblog.com/important-justifiable-and-constrained-role-nationality-foreign-intelligence-surveillance-0.

61. Christopher Bing and Joel Schectman, "Inside the UAE's secret hacking team of American mercenaries," *Reuters*, January 30, 2019, https://www.reuters.com/investigates/special-report/usa-spying-raven/.

62. Kovacs and Hawtin," Cybersecurity, surveillance and Online Human Rights"

63. Oona Hathaway and Scott Shapiro, " What realists don't understand about law," *Foreign Policy,* December 9,2017, https://foreignpolicy.com/2017/10/09/what-realists-dont-understand-about-law/.

64. Oona Hathaway and Scott Shapiro," International law and its transformation through the outlawry of war," *International Affairs* Volume 95 no 1 (2019):63.

65. Stephen Walt," There's still no reason to think the Kellogg-Briand Pact accomplished anything," *Foreign Policy*, September 29,2017, https://foreignpolicy.com/2017/09/29/theres-still-no-reason-to-think-the-kellogg-briand-pact-accomplished-anything/.

66. Oona and Shapiro, " What realists don't understand about law".

67. "Pegasus project: UN Experts call for moratorium on sale of surveillance technology," *The Wire*, August 12,2021, https://thewire.in/rights/pegasus-project-un-experts-moratorium-surveillance-technology

68. Hathaway and Shapiro, "International law and its transformation through the outlawry of war".

69. Arindrajit Basu,"The enterpreneurs of international law: A brief history," *RSRR Blog,* February 2, 2021, http://rsrr.in/2021/02/02/entrepreneurs-of-international-law/.

70. Jef Huysmans, " Democratic curiosity in times of surveillance," *European Journal of International Security* Volume 1 No 1 (2016):91.

71. Huysmans, " Democratic curiosity in times of surveillance".

Space-Nuclear Nexus: The Interface Between Key Technologies

Victoria Samson

From the very beginning, the space age has been interwoven with nuclear and strategic considerations. Part of this stems from its start as a competition of geopolitical dominance between the Cold War superpowers; part of this comes from a lot of the same technologies used in the nuclear strategic context (i.e., intercontinental ballistic missiles used to deliver nuclear warheads being repurposed to be used to launch satellites); part of this comes from the fact that many of the initial nuclear powers also being the early adopters of space capabilities. As nations' use of space continues to evolve, this relationship between nuclear and space technologies evolves as well, and the consequences of an unstable space domain can have an increasingly more costly result.

Cold War Considerations of Space and Nuclear

When considering the connection between space and nuclear issues, particularly when examining their relationship in a strategic context, often the first thought is of national technical means (NTMs) and noninterference thereof. NTMs are satellites that gather intelligence and are usually interwoven deeply into the nuclear strategic calculus. Since they are often used as part of the decision-making process in terms of ascertaining whether one of the major superpowers is under nuclear attack, they have been perceived as being untouchable (in a figurative sense) because their loss could lead to inadvertent escalation. Due to this, clauses that discourage interfering with the NTMs have been a part of strategic control agreements dating back to the 1972 Anti-Ballistic Missile Treaty. It is even part of the treaty between the US and Russia called the Measures for the Further Reduction and Limitation of Strategic Offensive Arms, aka New START (and the extended version thereof)[1].

Another link between space and nuclear is the early warning satellites intended to detect ballistic missile launches. For the US, that includes the Space-based Infrared System programme, which encompasses four satellites in geosynchronous Earth orbit (GSO) and two hosted payloads in highly elliptical orbits (HEO). A successor programme, Next Generation Overhead Persistent Infrared, is anticipated to be launched in 2025 with three satellites in GSO and two in HEO[2]. For Russia, this includes a small constellation of four new-generation missile early warning satellites called Tundra that fly in highly elliptical orbits. The Tundra satellites are part of the Integrated Space System, which will eventually also include several satellites in geostationary Earth orbit (GEO)[3].

Enter: Counterspace Capabilities

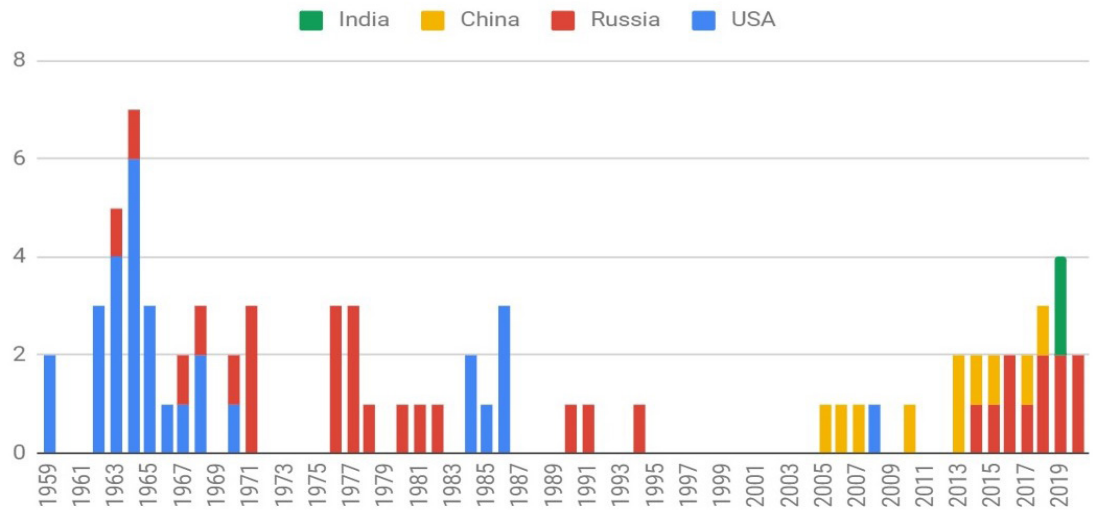
But there is a larger conversation to be had about the overall stability of the space environment having ripple effects on strategic relationships and nuclear stability. If there is debris created by an anti-satellite (ASAT) test, whether your satellite is related to strategic security or not, it will be threatened by it. If there are norms evolving that allow for non-cooperative close approaches to

other countries' satellites, your satellites could potentially be threatened by that. If you have a complicated space traffic management picture with tens of thousands of satellites cluttering orbits, this can have an effect on your ability to access and use satellites needed for strategic security. If there is hostility on Earth, this can lend itself to potential miscalculations, misunderstandings, and mistrust that could lead to unintended escalation in space.

And right now, one is seeing an upward trend in the proliferation of counterspace capabilities globally. An increasing number of actors (nation-states but also commercial entities and academic institutions) are becoming involved in space. While this eventually leads to benefits on Earth, it has the negative externality of resulting in a concomitant amount of competition and congestion in space. As more countries incorporate space capabilities into their national infrastructures and use it to enhance their military capabilities and national security, this has also led more countries to look at developing their own counterspace capabilities that can be used to deceive, disrupt, deny, degrade, or destroy space systems. There are global repercussions from their use that go well beyond the military sphere, as much of how our society is shaped and the global economic infrastructure is built on space applications.

Part of this proliferation is in the number of countries building up military space organisations as a way to protect their space assets. Russia and China both have military space organisations, and in 2019 alone, the US, France, Japan, and India created such originations (although it should be noted that the US's Space Force was built upon existing military bureaucracy). In the past year, Australia, Germany, and the UK have launched their own versions as well. While the existence of a military space organisation does not automatically translate to an offensive counterspace capability or interest, it does indicate that the host countries are evolving their use of space to include military aspects, and are open to considering developing some sort of counterspace capability (the reason given at least partially for why such a bureaucracy is needed is that there are threats to space capabilities and a response is needed).

Figure 1: ASAT Tests by Country and Year



Source: History of ASAT Tests in Space, Secure World Foundation, 2021[4]

Another part of this proliferation can be seen in a recent uptick in ASAT tests. Since 2005, there have been more than 20 tests of anti-satellite weapon systems in space by four different countries, a rate of testing that has not happened since the 1960s.

In addition to outright weapons testing, another concerning development is that some satellites have been observed to make deliberately close approaches to the satellites of other countries without coordination, prior knowledge, or consent of the operators of those satellites. There is growing concern that such behaviours might increase tensions between countries or be misinterpreted as a hostile action that precipitates an armed attack. Satellites related to nuclear command and control, if approached in a non-consensual manner by satellites of a geopolitical rival, could be perceived as being threatened and the satellites' owner may respond accordingly.

The Secure World Foundation's *Global Counterspace Capabilities*[5] analyses counterspace capabilities being developed by multiple countries across five categories: direct-ascent, co-orbital, electronic warfare, directed energy, and cyber. Direct ascent weapons are those that use ground, air-, or sea-launched missiles with interceptors to kinetically destroy satellites through force of impact but are not placed into orbit themselves. Co-orbital weapons are that those that are placed into orbit and then maneuver to approach the target to attack it by various means, including destructive and non-destructive. Directed energy weapons use focused energy, such as laser, particle, or microwave beams to interfere with or destroy space systems. Electronic warfare weapons use radiofrequency energy to interfere with or jam the communications to or from satellites. Finally, cyber weapons use software and network techniques to compromise, control, interfere, or destroy computer systems. While it is a complicated discussion, to generalise the current state of development globally, it should be noted that although there is significant research and development on a broad range of destructive and non-destructive counterspace capabilities in multiple countries, only non-destructive capabilities are actively being used in current military conflicts.

Finally, one more trend emerging is the development and expansion of space situational awareness (SSA) capacities, both at the national and the commercial level. SSA can be generally defined as information about the space environment, including on active satellites, space debris, and (in theory) activities on orbit. As such, reliable and accurate SSA is very important in verifying actions on orbit. While SSA is not uniquely used for counterspace, it is a critical enabler for both offensive and defensive counterspace operations. Furthermore, it is very important in verifying that agreements or guardrails to irresponsible behaviour are being followed, or, depending on one's perspective, to be able to identify when irresponsible or unusual or threatening behaviour on orbit is happening.

New Uses of Space and Effects on Strategic Relationships

The intersection of space and nuclear is not static; as the use of space evolves, so does the strategic relationship with nuclear. As new uses of space emerge, they challenge the current state of space governance and stretch it, often well beyond what the existing norms of behaviour and/or legal regime can comfortably support. Activities like active debris removal or on-orbit

servicing, assembly, and manufacturing (OSAM) can complicate the space environment to the possible point of instability, particularly if the actions are being undertaken by geopolitical rivals, which could then have consequences for the nuclear strategic relationship.

And returning to non-interference with NTMs, there are some in the US national security establishment who are worried about possible sneak attacks coming from cislunar space. Several of the US's most important and expensive national security satellites are in GEO, which is at an altitude of roughly 36,000 kilometers. It takes about seven hours for an object launched from Earth to reach GEO, so it is challenging to surprise satellites at GEO (unless the offensive satellite is already in GEO and undertakes a non-consensual close approach). Coming from the Moon, it will take about three days, but it requires much less energy to move an object from cislunar orbit to GEO than from the Earth to GEO (most of the energy required to get the Moon is used to get out of the Earth's gravity well). Accordingly, some of the more hawkish members of the US national security space community worry that an unfriendly country like China will position a satellite in cislunar orbit and then secretly alter its orbit to attack the US's NTMs in GEO.

What makes this scenario theoretically possible is that the US military's existing SSA capabilities are focused largely on objects in Earth orbit only. The Department of Defense does not currently have the ability to consistently track, catalogue, and monitor objects in cislunar space. While China has not expressed any interest or undertaken activities on or around the Moon to give it this type of capability, there is a small but vocal group within the US hyping the potential of that threat. Independent of this concern, there are other less-inflammatory reasons to have extra-GEO or cislunar situational awareness; with more countries striving for the Moon and more activities are happening on or around the Moon, it makes sense to enhance situational awareness for spaceflight safety at the very least.

Multilateral Responses to Space and Nuclear Connection

Space and nuclear have long been interwoven in terms of strategic arms control efforts. Early in the Cold War, the nuclear weapon tested in space as part of the Starfish Prime test in 1962 led to the Limited Test Ban Treaty, which prevented the signatories from testing nuclear weapons above ground or on orbit, due to concerns about the electromagnetic pulse released as part of the 1.45 MT test[6]. The connection between satellites being used as NTMs for strategic arms control treaties has already been mentioned. Perhaps even more elucidating is the way in which concerns about nuclear issues have resulted in the stagnation of multilateral discussions on space security issues.

Multilateral discussions have been stalled for decades on space security issues. The Conference on Disarmament is a consensus-driven institution that focuses on security-related issues. As part of its focus on consensus, the agenda for negotiations must be agreed to unanimously. Space security negotiations have not moved forward because the Conference cannot come to agreement about unrelated issues—in this case, fissile material. Another case of space and nuclear concerns shaping one another.

That is not to say that there have not been efforts to move ahead with space security-related discussions at the multilateral level; it is just that up until recently, there has been very little progress[7]. Much of the focus has been on the proposed draft 'Treaty on the Prevention of the Placement of Weapons in Outer Space, the Threat or Use of Force against Outer Space Objects' put forward by Russia and China in 2008 (and modified in 2014)[8]. This treaty has been criticised by the US[9] and its allies as not having verification mechanisms and also allowing for ground-based ASAT tests, and it has languished in multilateral discussions. Likewise, another measure promoted by Russia and China, 'No First Placement of Weapons in Space' (NFP), has found little support in international discussions, with the criticism by the US and its allies that NFP implies that weapons in space are fine, as long as they are used in response to someone else placing them in orbit first.

Exacerbating this is that there is a lack of agreement, broadly speaking, about what is the biggest threat to space security and stability and how to approach that threat. If the international community cannot come to agreement as to what the problem is, it is unlikely to make much headway in finding a solution. Is it the placement of dedicated weapons on orbit (namely, space-based missile defense interceptors), as Russia, China, and their allies believe? Or is it largely one that is almost environmental in nature—congestion and competition over increasingly complicated orbits and new uses of space—as the US and its allies posit? Further roiling negotiations is a disconnect between countries on what shape the solution should take. Should it be a legally binding treaty? Should it be agreed-upon norms of behaviour? Once again, Russia, China, and their allies believe in the former, while the US and its allies believe it should be the latter. This also ties into how various factions approach establishing space security and stability: should it be by banning threatening technologies, i.e., designated weapons, as Russia, China, and their supporters argue? Or due to the inherent dual-use nature of space, it is challenging to determine what technology is offensive in nature and so discussions should focus on behaviour in orbit, as the US and its supporters think?

With this complicated diplomatic setting, it is perhaps not surprising that there has been little movement toward collective action for enhancing space security and stability. However, over the past year, there has been some cause for cautious optimism. In December 2020, the UK co-sponsored a resolution in the United Nations General Assembly (UNGA), UNGA 75/36[10], which strived to reset some of the conversations on space security and stability. It ended up passing with broad global support. It asked nations to submit reports to the UN Secretary-General by May 2021 detailing three things—what they believe to be the most concerning threats to space security and stability; what they deem to be responsible (or irresponsible) behaviour in space; and finally, what they perceive to be the best way forward for international discussions on these issues. Over two dozen countries sent in submissions, plus a handful of civil society actors (including the Secure World Foundation).

In reading through the reports and the final report by the UN Secretary-General, there appear to be several norms emerging[11]. The first is that it is considered a bad idea to deliberately create long-lived debris on orbit. The second is that it is thought to be poor form to conduct non-consensual close approaches to other countries' satellites. (These two norms also show up in the US Secretary of Defense Austin's July 2021 memo listing five norms of behaviour in space that he wanted the Department of Defense to explore more fully[12]. The third is that there generally

seemed to be support for discussions to be held later at some sort of open-ended working group (OEWG), a very UN type of organisation that is not open-ended in terms of its time of work but in that any country can join the discussion (as opposed to a group of governmental experts, or GGE, which is generally limited to 15 countries, five of which are the permanent members of the UN Security Council). The UK has since indicated that it intends to submit another resolution to the UNGA this fall to create an OEWG that would strive to create international norms for responsible behaviour in space[13]. Perhaps some momentum will be created by this that will allow for some type of progress on space security and stability discussions that will then have positive effects on nuclear stability.

Conclusion

Space and nuclear issues have been connected for as long as there have been satellites. At first, this connection was seen in how satellites helped improve nuclear strategy stability. But as the use of space evolved, so did this relationship. Space security can very much affect strategic stability and thus shoring up the former can help improve the latter. The proliferation of counterspace capabilities, SSA radars and data, and even cislunar competition can all have strong effects on strategic stability. Hence it is crucial to continue to fully support good faith negotiations and international efforts to strengthen space security and stability. This in turn can help ensure a stable, predictable strategic environment for all.

Endnotes

1. US Department of State, Government of the United States of America, https://www.state.gov/new-start/, 2021.

2. Nathan Strout, "Space Force, Lockheed are ready to start making the nation's new satellites to watch for missiles," *C4ISRNet,* August 24, 2021, https://www.c4isrnet.com/smr/space-competition/2021/08/24/space-forces-next-generation-of-missile-warning-satellites-passes-major-design-milestone/

3. Bart Hendrickx, "EKS: Russia's space-based missile early warning system," *TheSpaceReview.com*, February 8, 2021, https://www.thespacereview.com/article/4121/1

4. "SWF Releases Updated Compilation of Anti-satellite Testing in Space," *Secure World Foundation*, June 30, 2020, https://swfound.org/news/all-news/2020/06/swf-releases-updated-compilation-of-anti-satellite-testing-in-space/

5. Brian Weeden and Victoria Samson, *Global Counterspace Capabilities: An Open Source Assessment*, Washington DC, Secure World Foundation, 2021, https://swfound.org/counterspace

6. Comprehensive Nuclear Test Ban Treaty Organization, *'Starfish Prime', Outer Space*, Vienna, Comprehensive Nuclear Test Ban Treaty Organization, 1962, https://www.ctbto.org/specials/testing-times/9-july-1962starfish-prime-outer-space

7. Victoria Samson and Brian Weeden, "Enhancing Space Security: Time for Legally Binding Measures," *Arms Control Today,* December 2020, https://www.armscontrol.org/act/2020-12/features/enhancing-space-security-time-legally-binding-measures

8. *Treaty on the Prevention of the Placement of Weapons in Outer Space, the Threat or Use of Force against Outer Space Objects*, 2014, https://reachingcriticalwill.org/images/documents/Disarmament-fora/cd/2014/documents/PPWT2014.pdf

9. U.S. Mission to International Organizations in Geneva, Government of the United States of America. https://geneva.usmission.gov/2019/08/14/statement-by-ambassador-wood-the-threats-posed-by-russia-and-china-to-security-of-the-outer-space-environment/. 2019.

10. *Reducing space threats through norms, rules and principles of responsible behaviours: Resolution / adopted by the General Assembly*, A/RES/75/36, Dec. 16, 2020, https://digitallibrary.un.org/record/3895440?ln=en

11. Office of the Secretary-General of the United Nations, *Report of the Secretary-General on reducing space threats through norms, rules and principles of responsible behaviors, 2021*, https://www.un.org/disarmament/topics/outerspace-sg-report-outer-space-2021/

12. Theresa Hitchens, "Exclusive: In A First, SecDef Pledges DoD To Space Norms," *BreakingDefense.com*, July 19, 2021, https://breakingdefense.com/2021/07/exclusive-in-a-first-secdef-pledges-dod-to-space-norms/

13. Theresa Hitchens, "Exclusive: UK Pushes New UN Accord On Military Space Norms," *BreakingDefense.com*, September 13, 2021, https://breakingdefense.com/2021/09/exclusive-uk-pushes-new-un-accord-on-military-space-norms/.

Norms

Understanding the UN Cyber Processes as Conflict Management Tools

Abagail Lawson

Malicious state action in cyberspace continues to increase, even as states reached a consensus on the norms for responsible behaviour through the UN Open-ended Working Group (OEWG) and Group of Governmental Experts (GGE) processes in 2021. Violations by states as well as continued uncertainty about how the agreed norms are meant to be implemented threaten to undermine the legitimacy and utility of the normative approach and cyber stability. These dynamics are fuelled by competing visions for the future of cyberspace, with some states seeking to assert an authoritarian version of state sovereignty in the domain, while others envision an open, interconnected future. This paper focuses on the current moment in the UN processes as part of an effort to manage conflict in cyberspace. It begins by describing the status quo in the domain, and then analyses three main elements of the UN discussions in light of concepts from conflict resolution literature to understand how the consensus reached in the OEWG and GGE can contribute to managing conflict in cyberspace, and where a fractured approach can be useful.

Discussions at the UN

The narrative around an international consensus on responsible behaviour in cyberspace has recently shifted. In 2019, the difficulty of the global conversation was on display when the UN General Assembly adopted two competing resolutions establishing separate processes on the use of information and communications technologies (ICTs) in the context of international peace and security—the OEWG and GGE. The GGE resolution, sponsored by the US, established the sixth iteration of a smaller group of countries coming together to discuss responsible behaviour in cyberspace. Previous GGEs issued reports laying out possible confidence-building measures, broad principles about the application of international law in cyberspace, and in 2015, articulated 11 voluntary norms for responsible state behaviour in cyberspace[1]. However, the fifth GGE was unable to agree on a consensus report in 2017, marking a period of division that carried into 2019, when Russia introduced the competing OEWG resolution. The OEWG was more inclusive, open to all member states of the General Assembly, and addressed many of the same topics as the GGEs, including the applicability of international law in cyberspace, current cyber threats, confidence building, capacity building and norms of behaviour.

The competing processes were chalked up to pessimism about the UN and evidence of deep rifts between major powers on cyber matters. But in March 2021, when the OEWG agreed on a report containing recommendations and affirming the conclusions of previous GGEs[2], the narrative reversed to celebrate the potential of multilateral diplomacy. This success was followed by the GGE consensus report in June 2021, which reinforced much of the OEWG's work and provided suggestions on how states might implement the 2015 norms, a seemingly herculean effort considering where the conversation had been in 2019[3].

Despite consensus on the reports, significant divisions remain between states, including on the issue of accountability for norm violations, as well as on what the goal of the international

negotiations should be. These divisions broadly reflect two competing models for cyberspace governance—a dynamic characterised as a "digital Cold War" due to the rivalry between authoritarian and democratic models championed by major powers, with a large group of countries caught in between[4]. The authoritarian governance model, notably backed by Russia and China, is based around state sovereignty in terms of non-interference in internal affairs, but also the ability of states to assert control of domestic online space. This group supported the OEWG and have called for negotiations on a legally binding treaty for cyberspace that could delineate sovereignty lines more clearly. The more democratic, "Western" model, supported by the US, EU and others, prioritises interconnectivity, openness, and principles like freedom of expression and access to information. These countries largely backed the GGE process and oppose a legally binding treaty, instead insisting that existing international law and voluntary norms are sufficient if applied and adhered to in cyberspace.

These divisions, notwithstanding the enthusiasm for the consensus reports, illustrate the reality of rulemaking for the cyber domain: hard and fast rules are difficult to come by as different states are pursuing different aims through these processes, and yet there is some level of acknowledgement that the status quo brings risks that need to be managed at an international level.

Status Quo: Cyber "Unpeace"

The situation in cyberspace has been characterised as a perpetual state of "unpeace" as states carry out malicious actions that fall below the threshold of armed conflict, but nevertheless have adverse, strategic cumulative effects[5]. Such state-sponsored actions are on the rise, as the number of significant incidents attributed to states more than doubled between 2017 and 2020, according to some threat researchers[6]. Several large-scale events have also highlighted the potential risks at stake: SolarWinds illustrated the vulnerability of supply chains, and the Colonial Pipeline ransomware incident raised concerns about critical infrastructure[7].

The current situation falls into a "grey zone" below the traditional threshold of armed conflict and yet is characterised by "conflict" in the non-legal sense of the word, referring to competition and struggle for power and advantage[8]. The ambiguity of this situation may have strategic advantages for states wishing to pursue their national interests and undermine adversaries without escalating to "armed conflict", which may have international legal implications and more serious downsides. However, the "unpeace" also undermines stability as it means unpredictability in how states interact with each other and how they may respond when faced with adversarial activity, as well as an overall lack of clarity about the rules of engagement in this increasingly militarised domain.

Conflict Management in Cyberspace

Despite the status quo, the level of engagement in the UN processes and recent consensus are positive indications that states, including major powers, recognise that the current instability poses a threat and should be managed. As noted in a discussion on confidence building measures (CBMs) for cyberspace, states have taken steps to engage in multilateral negotiations in part because they have recognised that "the secretive nature of cyber operations and the difficulties

of signalling in cyberspace can be destabilising to interstate relations, increasing tensions and the risk of inadvertent conflict"[9]. This recognition is also heard in the statements made during the OEWG, as many developing countries and states affiliating themselves with the Non-Aligned Movement decried the growing militarisation and weaponisation of cyberspace as a threat to other benefits from the use of ICTs, such as development gains from digitalisation, profits of tech companies subject to cyberattacks, and general trust in cyberspace a means for communication and business[10].

At the same time, if states believe there is strategic advantage to be gained by pursuing cyberattacks and find the resulting instability tolerable, expectations for what can be accomplished at the international level will be limited. The goal may not be to *eliminate* conflict in cyberspace, but instead we can understand efforts at the UN as part of a larger endeavour to *manage* the level of conflict in the domain, and as such, draw on conflict management literature, concepts and practices to advance the current international discussions.

The study of conflict management and resolution draws on sources from a wide range of disciplines, including anthropology, psychology, economics, peace studies, mathematics, law, and political science. The concepts are applied and practiced in a variety of settings, including corporate and industrial workplaces, judicial proceedings and criminal justice, grassroots justice movements and community activism, interpersonal mediation, and international diplomacy. The roles of various parties, models, tactics for negotiation, and principles for peacemaking and conflict resolution from the interstate to community level have been applied to international relations[11]. The sections below look specifically at three major aspects of the UN processes— CBMs, capacity building, and voluntary norms of behaviour[12]. While these concepts have been examined in their specific application to cyberspace[13], here they are mapped to three elements of conflict resolution: fostering genuine dialogue, democratising power, and establishing rules for "fair fighting". This analysis demonstrates that the UN cyber processes have value as mechanisms in a conflict resolution process, though the situation is not yet ripe for more definitive breakthroughs.

Foster genuine dialogue through CBMs

The role that dialogue can play in conflict resolution varies and depends on the actors, setting, and how dialogue processes are used. Dialogues can be focused on outcomes (concrete agreements) or on process, laying the groundwork for future agreement or simply building trust between the parties[14]. In intercultural dialogues, for example, dialogue is seen as a communicative interaction in which the parties come to understand themselves in relationship to others, as a necessary step in defusing tensions and preventing or resolving conflicts[15]. William Ury, founder of the Harvard Negotiation Project, writes about fostering dialogue as a core part of conflict resolution, even between opposing sides who lack agreement on anything of substance[16]. Dialogue does not need to aim at changing opinions or even reaching agreement on the issues. Its value can be in providing space for parties to talk, promoting mutual understanding, and building relationships to prevent escalation.

Both the OEWG and GGE are understood as dialogues that serve process-oriented purposes of communication and building understanding among states about their respective positions. Additionally, the CBMs proposed in these processes are particularly relevant as mechanisms of dialogue. CBMs such as sharing information on national policies or views on how international

law applies in cyberspace, can help increase transparency and the understanding of each other's postures, which helps interpret actions and prevent unintentional escalation in the "unpeace" environment. Other CBMs can contribute to facilitating direct communication in crises or after incidents, such as the establishment of points of contact within governments on the technical, diplomatic and political levels, which has been suggested within the GGE and OEWG, and established in regional contexts like the Organization for Security and Co-operation in Europe.

Democratising power through capacity building

Power plays a determinative role in conflict and efforts to resolve it. Peacemaking can be understood as achieving a balance of powers, "an interlocking of mutual interests, capabilities, and wills" [17]. Power dynamics within negotiations can also impact the outcomes and sustainability of the consensus achieved, and processes with asymmetrical dynamics can impede the likelihood of success[18]. While there are different ways of understanding "power" in international affairs, resources and the ability to affect outcomes are principal properties[19].

In the UN cyber processes, the power dynamics are highly asymmetrical. Cyber capacity among countries varies widely, both in terms of technical capabilities and infrastructure, and with regard to resources and structures for policy and governance. Not only does this change the way countries perceive threats, but it also impacts their ability to participate in international discussions and CBMs, and to implement norms. For example, the level of maturity of a country's computer emergency response team (CERT) will determine their ability to participate in CBMs focused on CERT-to-CERT cooperation. An immature national CERT with two staff members and no profile within the national security and political apparatus will not be an effective point of contact for exchanging information or cooperating across borders in the face of a large-scale incident. Similarly, if a country does not have an established procedure for dealing with cyber incidents, then nominating a point of contact for a directory will be less useful if the designated contact does not have the expertise or resources to ensure the flow of information to the right points within their government.

Cyber capacity building is a necessary step to enable countries to set their own policies for the use of cyber tools, and to participate in the rule-setting discussions at the international level. While many countries have taken steps to integrate cyber expertise into their foreign affairs ministries to facilitate international cooperation, many others lack such expertise or designated resources that makes their participation in global, bilateral and regional forums on these topics less effective.

Capacity building also raises overall levels of cybersecurity and resilience, strengthening a country's defences against malicious acts. This can also contribute to reduced tensions because a well-defended resilient country may be less worried about a certain level of threat. Underscoring the importance of cyber capacity building to enable implementation of norms and CBMs, the OEWG report outlines a series of principles to guide international capacity building efforts, including basing efforts on evidence, creating mutual trust between partners, and ensuring that efforts are inclusive and respect state sovereignty and human rights.

Establish rules for "fair fighting" through norms

In the conflict resolution field, it is acknowledged that conflict is not always avoidable and, in some cases, may be necessary for parties to air their grievances and bring about change[20]. In these cases, practitioners often focus on reducing harm and containing the conflict to prevent unnecessary violence and devastation. One approach to this is to establish rules for "fair fighting." Sometimes this involves getting parties to agree on what kinds of weapons are acceptable for use—examples from community violence prevention include getting gangs to settle disputes without guns and agreeing to a "fists only" rule. In couples counselling, "fair fighting" rules like dealing with one issue at a time or "no name calling" are commonly used approaches to manage conflict and each party's response when it arises or escalates. In the context of interstate conflict, the doctrine of mutually assured destruction at its most basic level established a "rule" for the US and Soviet Union—do not use nuclear weapons in a first strike.

The 2015 GGE norms are an effort to establish rules for fair fighting as states continue to use cyberspace as a strategic domain. Some of these norms are restraining on state action, such as designating critical infrastructure or CERTs as off limits to attack; others are prescriptive, describing responsibilities states should take, such as not allowing malicious ICT activity to be conducted on their territory. For instance, the 2021 ransomware incidents attributed to Russian criminals are being discussed under this latter norm. In this case, the US has argued that Russia has an obligation to act against the DarkSide criminal group in the wake of the Colonial Pipeline ransomware incident[21].

Limitations: Accountability and Ripeness

While dialogue, equalising power dynamics and establishing rules of the road are constructive additions to the goal of greater stability in cyberspace, there are limitations to what the UN processes can achieve. One of the divisions in the UN discussions is around accountability for violating the agreed-upon norms. Often, conflict resolution and mediation approaches use a third party to point out violations, a role likened to a referee[22]. However there is no referee in cyberspace and the enforcement of international law relies on the will of states. While proposals to establish international accountability mechanisms abound, they face steep political and practical hurdles[23]. For some, this signals the limitations of a norms-based, legalistic approach, and has implications for the foreign policy of individual states[24]. From a conflict resolution standpoint, this may indicate that that the situation is not yet ripe for deep consensus and norm adherence.

Ripeness is a key concept in conflict resolution and mediation, describing the characteristics of a situation that drive parties to agreement, usually when the situation reaches either a plateau (a mutually hurting stalemate in which all parties feel uncomfortable in the current situation without an end in sight) or a precipice (a realisation that the situation is rapidly becoming worse, with an impending crisis if it is unresolved and policy options that are usually limited to major escalatory actions with uncertain outcomes)[25]. The current situation in cyberspace is worsening in terms of more frequent, larger-scale attacks and many states have expressed concern. It is clear, however, that the current costs of malicious action are not seen as intolerable and neither a plateau nor precipice has been reached for major powers. The UN processes can help raise the voices of other countries concerned about the situation but will likely not existentially change the calculations for larger states until other costs are felt.

Conclusion

The status of international consensus on rules of the road in cyberspace neither indicates a failure of international diplomacy or a sure sign of success. While congratulating themselves upon the adoption of the OEWG report, many countries also expressed displeasure with significant aspects of the report. However, the OEWG and GGE outcomes represent a "line in the sand," according to one diplomat, from where the proverbial trudge up the beach can begin[26].

The next step is again going to be about process. Russia proposed a new OEWG before the first one finished, which held its first meeting in June 2021 and is underway until 2025. A Programme of Action was also proposed by the EU and 40 states as a mechanism to advance discussions about responsible state behaviour in cyberspace under one multifaceted process[27]. This fractured approach highlights the difficulties that remain but does not necessarily undermine the goal of international consensus. A fractured approach can also offer opportunities. The proliferation of initiatives for like-minded groups, such as the Paris Call for Trust and Stability in Cyberspace, the joint statement by 28 states on upholding the normative framework, and private sector-led initiatives like the Charter of Trust and the Cybersecurity Tech Accord are critical in socialising the norms and CBMs, building a body of thought around them, and generating new ideas and approaches towards their implementation. Like-minded groups focused on pointing out norm violations are also part of the process, such as when the US and allies issued a joint statement attributing the 2021 exploit of vulnerabilities in the Microsoft Exchange server to China's Ministry of State Security[28]. The current lack of interstate consensus on violations should not be seen as indicative of failure. Scholarship on international norm evolution indicates that when new norms are raised, there is a period of conflict between advocates of the new and supporters of the old[29]. As countries continue to affirm the 2015 GGE norms in public, multistakeholder processes like the OEWG, the hypocrisy of their violations are not without some cost and the "civilising force of hypocrisy" can eventually prompt shifts in behaviour[30].

Breakthroughs in conflicts often happen through shifts that were unimaginable at the start, and the UN processes represent nascent stages of resolution[31]. At this point, many states still see the status quo of "unpeace" as useful to achieve their strategic aims. The conflict management mechanisms to establish dialogue, build trust and mutual knowledge about national positions and understandings, and discuss red lines in cyberspace are crucial prerequisites for states to engage in deeper negotiations and possible resolution when the time is ripe.

Endnotes

1. United Nations General Assembly, *Report of the Group of Governmental Experts on Developments in the Field of Information and Telecommunications in the Context of International Security*, A/70/174 (July 22, 2015), https://undocs.org/A/70/174.

2. United Nations General Assembly, *Final Substantive Report of the Open-ended working group on developments in the field of information and telecommunications in the context of international security* (March 10, 2021), https://front.un-arm.org/wp-content/uploads/2021/03/Final-report-A-AC.290-2021-CRP.2.pdf.

3. United Nations General Assembly, *Report of the Group of Governmental Experts on Advancing Responsible State Behaviour in Cyberspace in the Context of International Security*, A/76/135 (July 14, 2021), https://front.un-arm.org/wp-content/uploads/2021/08/A_76_135-2104030E-1.pdf.

4. Inside Cyber Diplomacy, "The AU and ICTs for Development," CSIS podcast. 47:13 min, May 21, 2021, https://www.csis.org/node/60973.

5. "Unpeace" was coined by Lucas Kello in his book *The Virtual Weapon and International Order* (New Haven: Yale University Press, 2017). For discussion of the cumulative effects that arise from cyber-attacks below the threshold of armed conflict see, Richard Harknett and Michael P. Fischerkeller, *Persistent Engagement, Agreed Competition, Cyberspace Interaction Dynamics, and Escalation*, Washington DC, Institute for Defense Analyses, 2018, https://www.ida.org/-/media/feature/publications/p/pe/persistent-engagement-agreed-competition-cyberspace-interaction-dynamics-and-escalation/d-9076.ashx.

6. Michael McGuire, *Nation States, Cyberconflict and the Web of Profit*, HP Development Company, L.P., 2018, https://threatresearch.ext.hp.com/web-of-profit-nation-state-report/. For an overview of publicly-known state sponsored incidents, see Council on Foreign Relations, "Cyber Operations Tracker," https://www.cfr.org/cyber-operations/.

7. David Uberti and Kim S. Nash, "SolarWinds Hack Forces Reckoning with Supply-Chain Security," *Wall Street Journal*, January 14, 2021, https://www.wsj.com/articles/solarwinds-hack-forces-reckoning-with-supply-chain-security-11610620200; Ellen Nakashima and Lori Aratani, "DHS to issue first cybersecurity regulations for pipelines after Colonial hack," *Washington Post*, May 25, 2021, https://www.washingtonpost.com/business/2021/05/25/colonial-hack-pipeline-dhs-cybersecurity/.

8. Lucas Kello, "Cyber legalism: why it fails and what to do about it," *Journal of Cybersecurity* 7, no. 1 (2021), https://academic.oup.com/cybersecurity/article/7/1/tyab014/6343244.

9. Erica D. Borghard and Shawn W. Lonergan, "Confidence Building Measures for the Cyber Domain," *Strategic Studies Quarterly* 12, no. 3 (Fall 2018), https://www.airuniversity.af.edu/Portals/10/SSQ/documents/Volume-12_Issue-3/Borghard-Lonergan.pdf: 16.

10. See for example, the Non-Aligned Movement, "NAM Working Paper for the Second Substantive Session of the Open-ended Working Group on developments in the Field of Information and Telecommunications in the Context of International Security" (submission to the OEWG, New York, April 2020), https://front.un-arm.org/wp-content/uploads/2020/04/nam-wp-to-the-oewg-final.pdf.

11. A good overview of the field and prevailing themes in relation to international conflict is provided by Louis Kriesberg, "Contemporary Conflict Resolution Approaches," in *Leashing the Dogs of War: Conflict Management in a Divided World,* ed. Chester A. Crocker, Fen Osler Hampson, and Pamela Aall (Washington DC: United States Institute of Peace Press, 2013), pp. 455–476.

12. These were the main topics of the 2019 OEWG agenda, along with international law, threats in cyberspace and regular institutional dialogue, and are also reflected as major sections of the GGE reports.

13. Jason Healey et al., *Confidence-Building Measures in Cyberspace,* Washington DC, The Atlantic Council, 2014, https://www.files.ethz.ch/isn/185487/Confidence-Building_Measures_in_Cyberspace.pdf. On norms in cyberspace, see for example Joseph S. Nye, *The Regime Complex for Managing Global Cyber Activities,* Global Commission on Internet Governance: Paper Series No. 1, Cambridge, MA, The Centre for International Governance, May 2014, https://www.belfercenter.org/sites/default/files/legacy/files/global-cyber-final-web.pdf.; and Paul Meyer, "Norms of Responsible State Behaviour in Cyberspace," in *The Ethics of Cybersecurity*, ed. Markus Christen, Bert Gordijn, Michele Loi, The International Library of Ethics, Law and Technology, vol 21 (Open Access: Springer, Cham, 2020), https://link.springer.com/chapter/10.1007%2F978-3-030-29053-5_18.

14. Pernille Rieker and Henrik Thune, *Dialogue and Conflict Resolution: Potential and Limits,* (London: Routledge, 2015).

15. Mike Hardy and Serena Hussain, "Dialogue in a Rapidly Changing World: Practitioner Assessments of the Potency of Intercultural Dialogue for Improving Social Cohesion," *Journal of Dialogue Studies* 7 (2019), http://www.dialoguestudies.org/wp-content/uploads/2019/12/Dialogue-in-a-Rapidly-Changing-World-Practitioner-Assessments-of-the-Potency-of-Intercultural-Dialogue-for-Improving-Social-Cohesion.pdf.

16. William Ury, *The Third Side: Why We Fight and How We Can Stop* (London: Penguin Books: 2000), pp. 135.

17. R.J. Rummel, "Chapter 10: Principles of Conflict Resolution," in *Understanding Conflict and War: The Just Peace: Vol. V* (Beverly Hills, CA: Sage Publications, 1981), https://www.hawaii.edu/powerkills/TJP.CHAP10.HTM#2.

18. Rieker and Thune, *Dialogue and Conflict Resolution: Potential and Limits*.

19. See Joseph S. Nye, Jr. *Soft Power: The Means to Success in World Politics* (New York: PublicAffairs Books, 2005).

20. Ury, *The Third Side: Why We Fight and How We Can Stop*, xix.

21. Joseph Marks, "Russia agrees to cyber rules and violates them at the same time," *The Washington Post*, July 14, 2021, https://www.washingtonpost.com/politics/2021/06/14/cybersecurity-202-russia-agrees-cyber-rules-violates-them-same-time/.

22. Ury, *The Third Side: Why We Fight and How We Can Stop,* pp. 179.

23. For some discussion of the challenge of accountability in the cyber domain, see Jacqueline Eggenschwiler, "Accountability Challenges Confronting Cyberspace Governance," *Internet Policy Review* 6, no. 3 (2017), https://policyreview.info/articles/analysis/accountability-challenges-confronting-cyberspace-governance.

24. Kello, "Cyber legalism: why it fails and what to do about it."

25. I. William Zartman and Saadia Touval, "International Mediation," in *Leashing the Dogs of War: Conflict Management in a Divided World,* ed. Chester A. Crocker, Fen Osler Hampson, and Pamela Aall (Washington DC: United States Institute of Peace Press, 2013), pp. 437–454.

26. Inside Cyber Diplomacy, "A Guide to the UN GGE," CSIS podcast, 50:46 min, June 11, 2021, https://www.csis.org/node/61229.

27. United Nations *Concept-note on the organizational aspects of a Programme of Action for advancing responsible State behaviour in cyberspace,*" December 2021, New York, UN, https://front.un-arm.org/wp-content/uploads/2020/12/sponsors-oewg-concept-note-final-12-2-2020.pdf.

28. The White House, Government of the United States of America, https://www.whitehouse.gov/briefing-room/statements-releases/2021/07/19/the-united-states-joined-by-allies-and-partners-attributes-malicious-cyber-activity-and-irresponsible-state-behavior-to-the-peoples-republic-of-china/.

29. Martha Finnemore and Kathryn Sikkink, "International Norm Dynamics and Political Change," *International Organization* 52, no. 6 (1998), pp 887-917, https://pos-graduacao.uepb.edu.br/ppgri/files/2016/02/Finnemore-and-Sikkink.-International-Norm-Dynamics-and-Political-Change-1.pdf.

30. Jon Elster, "Deliberation and Constitution Making," in *Deliberative Democracy*, ed. Jon Elster (Cambridge: Cambridge University Press, 1998), 97–122.

31. Nelson Mandela, "Time and again conflicts are resolved through shifts that were unimaginable at the start," (speech, Dublin, April 13, 2000), The Independent, https://www.independent.co.uk/news/world/europe/independent-lecture-delivered-by-nelson-mandela-279789.html

Codes and Coalitions: Path to Global Governance of Artificial Intelligence?

Smriti Parsheera

Technology governance approaches are often cast as binaries—democratic or authoritarian, liberal or illiberal, and innovation-enhancing or rights-preserving. A closer look at the actions of state actors, however, suggests that domestic responses to critical technologies are far from settled. Similarly, governance by non-state actors also remains fluid and erratic. This is often because the boundaries of critical technologies and the scope of their benefits and risks remain contested. Current debates around the governance of artificial intelligence (AI) demonstrate many of these characteristics. Several voluntary codes of conduct and national strategies have come up to try and address the social, economic, ethical, environmental, and governance implications of AI. But the emerging nature of such technologies and their diverse applications and impacts tends to make AI governance a moving target.

Against the background of evolving use cases, national priorities, and polarised international classifications, a global consensus on the governance of critical technologies seems far from sight. In the interim, countries have started turning to multilateral, minilateral and bilateral coalitions to shape their cooperation on, and the governance of, critical and emerging technologies. The Organisation for Economic Co-operation and Development (OECD) principles on AI, the Global Partnership on Artificial Intelligence (GPAI), the Quad Critical and Emerging Technology Working Group, and the D20 are some examples[1].

Which Technologies, What Forms of Governance?

To understand the context and scope of international coalitions on emerging technologies, it would be useful to begin with a brief conceptual framing on two questions. First, what are the types of technologies that are being classified as emerging or critical, and to what extent do these categories coincide? Second, what are the governance modalities that become relevant in this context?

In July 2020, the Australian government set up a Critical Technologies Policy Coordination Office to guide its strategy on the developments, opportunities and risks related to critical technologies. They define such technologies to mean "current and emerging technologies with the capacity to significantly enhance, or pose risk to, our national interests (economic prosperity, social cohesion and/or national security)"[2]. The criticality of the technology in this case, therefore, stems from its relationship with national interests, which itself is defined fairly broadly. The Quad alliance, which has Australia, India, Japan and the US as its members, has opted for an even broader framing. They include both emerging and critical technologies within the ambit of a newly formed working group on these subjects[3].

The general understanding of emerging technologies accounts for characteristics such as novelty, fast pace of growth, socioeconomic implications, and a certain degree of uncertainty and ambiguity[4]. Commonly cited examples of such technologies include AI, 5th generation

(5G) telecommunications networks, blockchain, Internet of Things, and quantum computing[5]. Policy documents and academic literature are replete with discussions about the transformative potential of these technologies but also the harms that they could cause. This includes hard impacts of technology on aspects like human rights, health, safety and environment, as well as softer impacts in spheres like social relations and morality, issues which find a smaller space in political debates[6].

This paper will focus on the governance of AI. The classification of AI as an emerging technology seems obvious and yet fallacious at the same time. AI as a field has been around for over six decades, making it a somewhat old technology. But recent advances in the scale, capabilities and significance of AI have spurred calls for a more meaningful governance. This has resulted in a proliferation of AI principles in national AI strategies, private sector codes, and several civil society and multistakeholder initiatives[7]. Intergovernmental cooperation on this issue, however, has remained limited. The OECD's AI principles, adopted in May 2019, were the first notable exception[8]. Since then, these principles have also been endorsed at the 2019 G20 summit at Osaka[9], and by the 19 member states that currently form part of the GPAI initiative[10].

Before getting into the specifics of these developments, it would also be useful to clarify the context in which 'governance' is being used in this paper. To avoid any conflation between governance and regulation, it is worth clarifying that regulation is only one type of governance. It usually involves the use of formal tools like laws, treaties, and conventions. But governance can also be informal, relying on soft law, networks, partnerships, and collaborations to shape the relationships between different actors[11]. Much of what we are currently seeing in the form of alliances between 'like-minded nations' on emerging and critical technologies falls under the domain of informal governance.

The concept of governance can therefore be seen through several lenses, including based on the actions involved, responsible actors, and the degree of formality in the process. Academics Jon Pierre and B. Guy Peters offer a useful breakup for four activities that collectively amount to the practice of governance—articulating a common set of priorities for society; coherence and coordination; capacity for steering; and accountability for actions[12]. They posit that government actors maintain a central, although not exclusive, role in the performance of these activities. Beyond this, each type of emerging technology has its own governance networks, consisting of a complex set of actors with varying interests and strategies[13]. Ultimately, it is the interaction between numerous state and non-state actors, like firms, consumers, civil society groups, and international organisations, that shapes technological governance. This understanding is supplemented by Lawrence Lessig's powerful categorisation of 'code as law' [14] and the literature on co-production of technology and society, which demonstrates how societal forces can shape the contours and trajectory of technology[15].

Emergence of International AI Coalitions

The internationalisation of the AI governance debate in the past few years can be traced back to the G7 Ministerial Declarations at Kagawa, Japan, in 2016 and Montreal, Canada, in 2018[16]. The latter contained a specific declaration of a vision for the development of 'human-centric' AI, which was further elaborated through the Charlevoix common vision document[17]. Soon after,

the OECD adopted its recommendations on AI in 2019, which consists of five principles and five recommendations on fulfilment through national policies and international cooperation. The five principles aimed at creating responsible stewardship of trustworthy AI are:

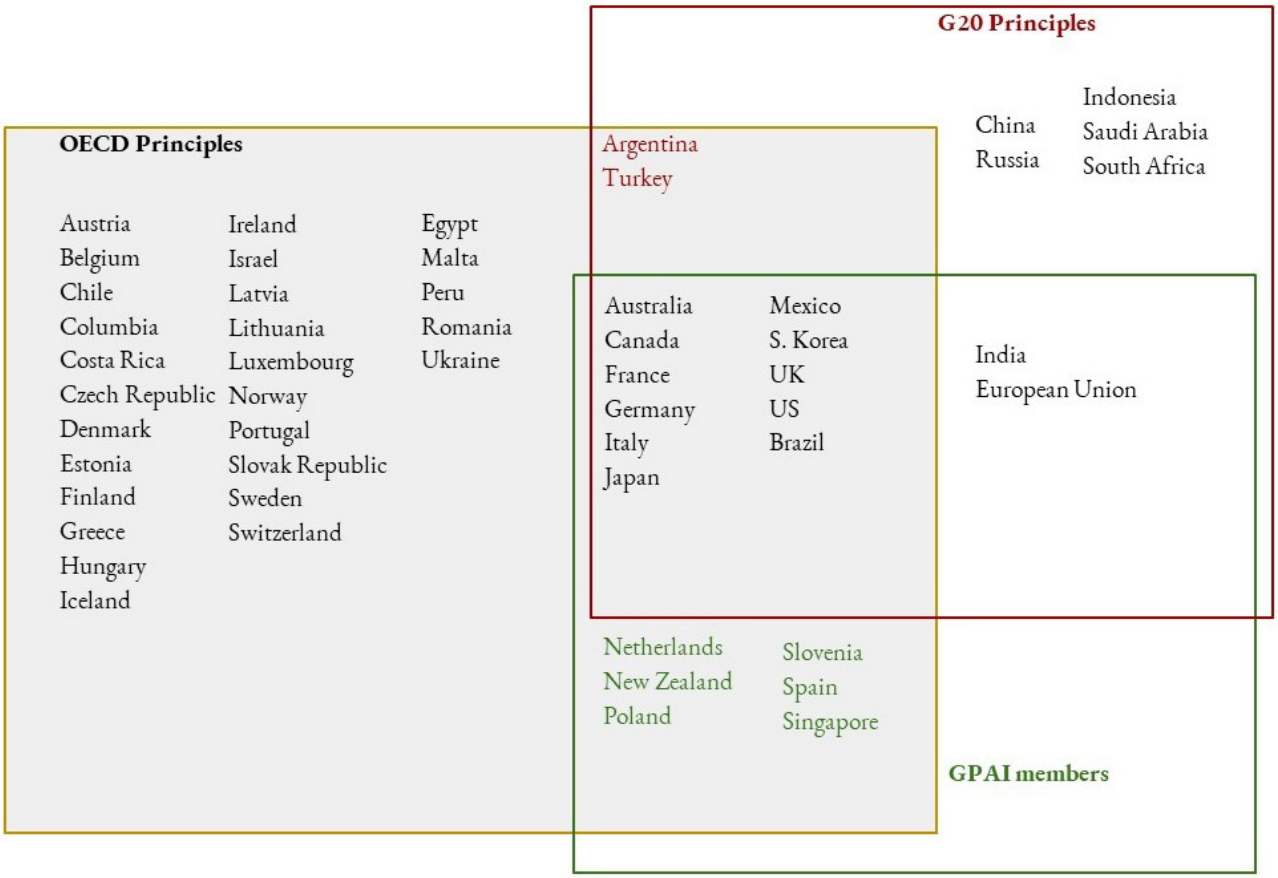
- **Inclusive growth, sustainable development and well-being** in terms of augmenting human capabilities, advancing inclusion, reducing inequalities, and protecting natural environments.
- **Human-centred values and fairness** through respect for the rule of law, human rights and democratic values.
- **Transparency and explainability** to foster a general understanding of AI systems and enable those affected by such systems to challenge their outcomes.
- **Robustness, security and safety** throughout the lifecycle of AI systems.
- **Accountability of AI actors** for the proper functioning of AI systems and for the respect of these principles.

Alongside these principles, the OECD recommendations urge adhering nations to invest in AI research and development, shape an enabling AI policy environment, build human capacity, and pursue international cooperation. This text finds clear parallels in the contents of the Montreal Declaration. Similar priorities have also been identified in numerous other national level and private AI principles. But what distinguishes the OECD conversation is the direct involvement of a fairly large number of countries in the deliberations and adoption—so far, the recommendations have been adopted by 38 OECD member states and eight non-members[18]. Although the principles are non-binding, the OECD envisages ongoing monitoring of the AI strategies adopted by adhering countries. The tasks of monitoring implementation, framing practical guidance, and promoting information sharing among stakeholders falls upon the OECD Committee on Digital Economy Policy acting through its AI Observatory.

While the OECD principles speak the language of human values, safety and accountability, AI governance also bears massive economic implications. This holds true for countries and businesses that are investing in these technologies. The attempts by the G7 and OECD, where many of these business interests are housed, in pursuing international discussions on AI are, therefore, unsurprising. But as noted earlier, the principals have also received broader support in forums like G20 and the GPAI. The three AI initiatives—G20 principles, OECD principles and GPAI—have common member countries, with the G7 states and a few others playing a central role in all discussions (see Figure 1).

The G20's support for the AI principles developed by OECD at the Osaka meeting was a notable development for a couple of reasons. The group represents 20 of the world's most systemically important advanced and emerging economies. The involvement of the BRICS countries, and other emerging economies like Indonesia and Saudi Arabia, therefore, added more global legitimacy to the OECD's AI agenda. Academics Jan Wouters and Sven Van Kerckhoven have illustrated how the OECD and G20 often serve as mutual reinforcements for each other's agendas; for instance, the OECD's Tax Information Exchange Agreements saw a massive surge in adoption due to political endorsement by the G20. The authors also highlight that this bonhomie can sometimes work to the disadvantage of non-member countries[19].

Figure 1: Country Participation in International AI Discussions



More specifically, the G20 is the only international conversation on AI principles that includes China as a participant. The country's interest in staying closely engaged in this conversation is evidenced by President Xi Jinping's comments at the 2020 G20 Summit: "China proposes a meeting on artificial intelligence (AI) in due course to advance the G20 AI Principles and set the course for the healthy development of AI globally"[20].

While discussing China's role, it is worth distinguishing between AI governance principles and the standardisation initiatives that are taking place in forums like the International Telecommunication Union (ITU). ITU's AI footprint includes standards on 'machine learning for 5G' and AI-enabled applications in areas like multimedia and data centre infrastructure[21]. The Chinese government and companies like Huawei are known to be fairly active in these debates[22]. Reports have also pointed to Chinese companies taking a lead in shaping ITU standards around facial recognition, video monitoring, and city surveillance[23]. These developments on the standardisation front are seen as one of the triggers for pushing forward the debate on AI principles and norms.

As one of the world's leading AI powers, China's absence from the GPAI, which is the first international multistakeholder initiative dedicated to AI, is both conspicuous and deliberate[24]. GPAI was launched in mid-2020 with 15 founding members[25]. Membership is now open to all countries, provided they endorse the values reflected in the OECD recommendations[26]. The various elements of GPAI's formal structure include a council, a steering committee, a secretariat,

a multistakeholder experts group plenary, and issue-based working groups. The GPAI council will consist only of member states but the steering committee will have a majority (six out of eleven) of non-government participants. In the spirit of multistakeholderism, this will include representatives from the fields of science, industry, civil society, labour/trade unions, and an international organisation.

The GPAI's terms of reference clarify that it will not develop norms or work on issues of national defence. However, non-adherence to the OECD principles could be grounds for termination of membership. This represents a form of semi-formal governance, with the OECD principles becoming the de facto standard of governance. The OECD's appointment as the GPAI secretariat further cements the linkages between the two forums. Further, two centers of expertise have been set up in France and Canada. Despite the global labeling, this institutional setup indicates a West-centric start for the GPAI.

Barrier or Path to Global Governance?

What impact will the various AI coalitions have on global AI governance? The term global governance "encompasses activity at the international, transnational, and regional levels, and refers to activities in the public and private sectors that transcend national boundaries[27]." By definition, this includes actors and actions that go much beyond the 54 countries and the European Union that are currently in this debate. Yet, all the forums previously discussed rely on a state-centric model for determining the pathways to AI governance. This is subject to some exceptions, as in the case of GPAI, which has tried to build in a more multistakeholder approach, although the model clearly remains a state-led one. Moreover, the participation of nation states in these debates remains far from equitable.

The current design of international AI discussions, therefore, poses certain barriers in terms of priorities, participation, and perspectives. This can present challenges in identifying the right goals and building enduring coalitions.

Priorities: The idea of human-centric and trustworthy AI lies as the heart of the OECD's AI principles. By extension, this is also mirrored in all the other initiatives. Many more countries are likely to endorse these principles in the years to come to gain membership into these exclusive clubs. But are these priorities that were originally set by a group of like-minded, democratic, largely market-driven economies, necessarily the most suitable? Jessica Cussins Newman points to how the focus on human centeredness is one of the main points of disconnection between the G7/OECD led and Chinese approaches. She refers to the work of Chinese scholars who emphasise the need for a more `harmonious' approach. This would account for the coexistence of human and AI alongside other goals like sustainable development and avoidance of monopolies[28]. Divorced from the labels of a Western-led or China-led approach, many countries, particularly in the developing world, might find themselves in sync with similar priorities.

Participation: Almost all the countries that are currently a part of AI discussions belong either to the group of advanced or emerging economies. The exclusion of less developed nations from these conversations leaves them to either accept the available principles as fait accompli or remain excluded from the gains of AI knowledge sharing systems. These concerns are compounded by the gap in the available evidence on the impact of new technologies in developing countries

and the limited public debate on technology in these regions[29]. At the same time, the role to be played by the OECD and the centers of expertise in France and Canada in the GPAI's activities could lead to questions of effective participation within the alliance. Any attempts at multistakeholderism within such forums will also have to contend with the risk of prioritising the participation of non-state actors from select countries on the grounds of relative 'expertise'.

Perspectives: Concepts like fairness, robustness, and accountability, while universally appealing, are not always easy to translate into practice. On the one hand, there is the concern that these principles might be stuck at remaining lofty goals without converting into tangible outcomes. On the other, the principles could also end up being applied in a manner that is not suited for the specific social, economic, and cultural context of each country[30]. The emphasis on practice guidance, best practice notes, and comparable metrics could add to the drowning out of differentiated perspectives.

The emphasis on 'like-mindedness' as the basis for the formation of the coalitions also stands to be questioned. This is true both in terms of the desirability of such an approach and whether such like-mindedness exists among the coalition partners. On the first point, the concept of like-mindedness often becomes a tool for the othering of countries like China and Russia, which will continue to pursue their own AI agendas outside of these conversations. While competition among AI principles could be a good sign in general, the clearly demarcated boundaries of 'us vs them' makes it unlikely to see countries switching between frameworks. An insistence on likeness therefore sets the stage for parallel and fragmented AI governance systems. At the same time, the partnering states are at different stages of the technological development and adoption cycle. They also display differing dynamics in terms of the effectiveness of courts and democratic institutions, state capacity, role of industry, and robustness of civil society engagement. Further, it is well recognised that technologies often foster a politics of their own[31]. The use of facial recognition technology for surveillance and law enforcement purposes is a case in point. While many countries have rightly called out China for its rampant use of facial recognition, particularly for surveillance against marginalised groups like the Uyghurs, this technology is being used by states across the political spectrum[32]. The use of facial recognition for surveillance could end up creating an authoritarian form of political life even in countries that are otherwise democratic in character, albeit with slightly better institutional safeguards[33]. Assumptions of likeness or otherness based on broad political or value systems may therefore be illusory.

Countries also differ in their approach to data governance, particularly on the free flow of data, which can have significant implications for AI governance. Alongside the adoption of the AI principles, the 2019 G20 meeting at Osaka also saw the initiation of a new initiative for discussions on 'data free flow with trust'[34]. India, Indonesia and South Africa, however, refused to go down the path of what is known as the Osaka Track. India's reservations stem from its desire to preserve policy space to make rules around data protection and e-commerce, and concerns that uninhibited cross border flows will hamper data access[35]. Among the GPAI members, India and Slovenia are the only countries that are not part of the Osaka Track[36]. While data flows are not explicitly mentioned in the AI principles, it may only be a matter of time before this discussion crops up in the context of creating an `enabling policy environment for AI'.

The limitations highlighted above, however, do not make a case for the abandonment of the non-binding coalition approach. Rather, they should be seen as inputs for refining—and recognising the limitations of—GPAI-like forums. As things stand, the risks posed by the unchecked growth of AI are real and immediate, but the possibility of a global accord through traditional forums like the UN is not. The roadblocks encountered in the work of the UN-constituted Group of Governmental Experts in the area of lethal autonomous weapon systems is illustrative in this regard[37].

In addition to being infeasible in the short to medium term, any sort of binding international norms may also not be undesirable at this point of time. One of the defining features of emerging technologies is the uncertainty surrounding their applications and outcomes. This implies that either the outcome of using the technology in a particular context or the probability of that outcome are not known in advance[38]. As a result, the complexity of choosing the right governance approach and the possibility of government failures is relatively high. Governance models therefore need to be tentative and adaptive in nature. This can be understood to mean an approach that is incremental, non-finalising and prudent in the sense that it relies on the lessons from a trial-and-error process[39]. Binding international commitments do not offer scope for such malleability.

Conclusion

Non-binding coalitions on emerging and critical technologies, such as the AI principles adopted by OECD members and endorsed by G20 and GPAI, appear to be the new normal in global technology governance. These frameworks could offer the benefit of peer learning and informal scrutiny without the heavy handedness of binding international norms, which appear both infeasible and undesirable at this time. But the coalition-based approach comes with its own set of challenges, particularly in terms of gaps in priorities, participation, and perspectives.

Except for the G20 forum, which represents a slightly broader spectrum of participants, international discussions on AI principles have so far been limited to a 'like-minded' set of countries. This has resulted in the proliferation of one set of principles developed by the G7/ OECD to multiple forums. But healthy competition among AI principles is as vital as competition among AI technologies themselves. Instead of insisting on adherence to the OECD AI recommendations, the G20 and GPAI should encourage participating states to use those principles as a starting point for further deliberations. That would prevent the imposition of the values of a few states on a vast majority of others, while allowing the injection of new ideas and priorities into the AI governance debate. The G20, which includes participation from China and Russia, presents a particularly important site for breaking away from the mould of like-mindedness toward more broad-based solutions.

Endnotes

1. Arindrajit Basu and Justin Sherman, "Two New Democratic Coalitions on 5G and AI Technologies," *Lawfare*, August 6, 2020, https://www.lawfareblog.com/two-new-democratic-coalitions-5g-and-ai-technologies; Mark Linscott and Anand Raghuraman, "How to Leverage the Quad to Counter China's Digital Sinosphere," *Atlantic Council*, May 17, 2021, https://www.atlanticcouncil.org/blogs/new-atlanticist/how-to-leverage-the-quad-to-counter-chinas-digital-sinosphere/.

2. Australian Government, Department of Prime Minister and Cabinet, *Protecting and Promoting Critical Technologies*, 2021, https://pmc.gov.au/sites/default/files/publications/protecting-and-promoting-critical-tech.pdf.

3. Ministry of External Affairs, Government of India, *Quad Summit Fact Sheet*, 2021, https://www.mea.gov.in/bilateral-documents.htm?dtl/33621/Quad_Summit_Fact_Sheet.

4. Daniele Rotolo, Diana Hicks and Ben R. Martin, "What is an Emerging Technology?," *Research Policy* 44, no. 10 (2015): 1827–1843, https://www.sciencedirect.com/science/article/abs/pii/S0048733315001031

5. G20 Ministerial Statement on Trade and Digital Economy, June 9, 2019, https://www.mofa.go.jp/files/000486596.pdf; Australia-India Cyber and Critical Technology Partnership – Grants Round 1 (2020), https://india.highcommission.gov.au/ndli/AICCTP.html.

6. Marianne Boenink, Tsjalling Swierstra and Dirk Stemerding, "Anticipating the Interaction Between Technology and Morality: A Scenario Study of Experimenting with Humans in Bionanotechnology," *Studies in Ethics, Law, and Technology* 4, no.2 (2010).

7. Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar, "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI", Berkman Klein Center for Internet & Society, 2020, http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420.

8. "Recommendation of the Council on Artificial Intelligence," OECD/LEGAL/0449, Adopted on 22 May 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449#adherents.

9. "G20 AI Principles," G20 Ministerial Statement on Trade and Digital Economy, 2019, https://www.g20-insights.org/wp-content/uploads/2019/07/G20-Japan-AI-Principles.pdf.

10. "The Global Partnership on Artificial Intelligence," https://www.gpai.ai/community/.

11. Asbjorn Roiseland, "Informal Governance," in *International Encyclopedia of Political Science*, ed. Bertrand Badie, Dirk Berg-Schlosser, and Leonardo Morlino, (Sage Publications, 2011), pp. 1020.

12. Jon Pierre and B. Guy Peters, *Governing Complex Societies: Trajectories and Scenarios*, (Palgrave Macmillan, 2005) pp. 3-6.

13. Eva Sorensen and Jacob Torfing, "Governance networks," in *International Encyclopedia of Political Science*, ed. Bertrand Badie, Dirk Berg-Schlosser and Leonardo Morlino, (Sage Publications, 2011), pp. 1030.

14. Lawrence Lessig, *Code Version 2.0.*, (Basic Books, 2006), http://codev2.cc/download+remix/Lessig-Codev2.pdf.

15. Sheila Jasanoff, "The Idiom of Co-Production," in *States of Knowledge: The Co-Production of Science and Social Order*, ed. Sheila Jasanoff, (Routledge, 2006), pp.1-12.

16. "Joint Declaration by G7 ICT Ministers (Action Plan on Implementing the Charter)," Takamatsu, Kagawa, Japan, April 30, 2016, http://www.g8.utoronto.ca/ict/2016-ict-declaration.html; "G7 Innovation Ministers' Statement on Artificial Intelligence," Montreal, Canada, March 28, 2018, http://www.g8.utoronto.ca/employment/2018-labour-annex-b-en.html.

17. "Charlevoix Common Vision for the Future of Artificial Intelligence," June 2018, https://www.international.gc.ca/world-monde/assets/pdfs/international_relations-relations_internationales/g7/2018-06-09-artificial-intelligence-artificielle-en.pdf.

18. Columbia and Costa Rica adopted the AI recommendations before becoming members of the OECD Convention.

19. Jan Wouters and Sven Van Kerckhoven, "The OECD and the G20: An Ever Closer Relationship?," Leuven Centre for Global Governance Studies, Working Paper No. 71, 2011 ,https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1898704.

20. Ministry of Foreign Affairs of the People's Republic of China, "Xi Jinping Attends Session I of 15th G20 Leaders' Summit and Delivers a Keynote Speech," November 21, 2020, https://www.fmprc.gov.cn/mfa_eng/zxxx_662805/t1834877.shtml.

21. "Interview with Chaesub Lee, International Standards for an AI-enabled Future," *ITU News*, July 6, 2020,

22. https://news.itu.int/international-standards-for-an-ai-enabled-future/,

23. Todd Shields and Alyza Sebenius, "Huawei's Clout Is So Strong It's Helping Shape Global 5G Rules," *Bloomberg Quint*, February 1, 2019, https://www.bloombergquint.com/global-economics/huawei-s-clout-is-so-strong-it-s-helping-shape-global-5g-rules.

24. Anna Gross, Madhumita Murgia, and Yuan Yang, "Chinese Tech Groups Shaping UN Facial Recognition Standards," *Financial Times*, December 1, 2019, https://www.ft.com/content/c3555a3c-0d3e-11ea-b2d6-9bf4d1957a67.

25. Joshua P. Meltzer and Cameron F. Kerry, *Strengthening International Cooperation on Artificial Intelligence*, Brookings, February 17, 2021, https://www.brookings.edu/research/strengthening-international-cooperation-on-artificial-intelligence/; Sushovan Sirkar, "India Joins US, EU in Global AI Group Seen As A Counter to China," *The Quint*, June 16, 2020, https://www.thequint.com/tech-and-auto/gpai-artificial-intelligence-india-joins-g7-nations-eu-in-policy-body-china#read-more.

26. The partnership now has 19 official members. Brazil, the Netherlands, Poland and Spain joined GPAI in December 2020.

27. "GPAI Terms of Reference," https://www.gpai.ai/about/gpai-terms-of-reference.pdf.

28. Kennette Benedict, "Global Governance," in *International Encyclopedia of the Social & Behavioral Sciences*, ed. Neil J. Smelser and Paul B. Baltes, (Elsevier, 2001).

29. Jessica Cussins Newman, "AI Principles in Context: Tensions and Opportunities for the United States and China," *Asia Society*, August 20, 2020, https://asiasociety.org/sites/default/files/inline-files/Cussins_Principles_Final.pdf.

30. Jonathan Wong, Tengfei Wang and Phadnalin Ngernlim, "Artificial Intelligence in Asia and the Pacific," *United Nations Economic and Social Commission for Asia and the Pacific*, November, 2017, https://www.unescap.org/sites/default/files/ESCAP_Artificial_Intelligence.pdf.

31. Nithya Sambasivan, Erin Arnesen, Ben Hutchinson, Tulsee Doshi, Vinodkumar Prabhakaran, "Re-Imagining Algorithmic Fairness in India and Beyond," Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021, https://doi.org/10.1145/3442188.3445896; Vidushi Marda and Shivangi Narayan, "On the Importance of Ethnographic Methods in AI Research," *Nature* Machine Intelligence, Vol 3, March 2021, 187–189

32. Langdon Winner, "Do artifacts have politics?, *Daedalus* 109, no. 1 (1980): 121–136, https://www.cc.gatech.edu/~beki/cs4001/Winner.pdf.

33. Steven Feldstein, *The Global Expansion of AI Surveillance*, Carnegie Endowment for International Peace, September17 2019, https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847.

34. Sarayu Natarajan and Smriti Parsheera, "A Conversation on AI Activism," Interactions 28, no. 1 (2021): 28–33, https://doi.org/10.1145/3436946.

35. Ministry of Foreign Affairs of Japan, *Osaka Declaration on Digital Economy*, 2019, https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/pdf/special_event/en/special_event_01.pdf.

36. Ministry of Commerce & Industry, Government of India, September 22, 2020, https://pib.gov.in/PressReleseDetail.aspx?PRID=1657874.

37. Arindrajit Basu, "Sovereignty in a 'datafied' world: A framework for Indian diplomacy," in *A 2030 Vision for India's Economic Diplomacy*, ed. Malancha Chakrabarty and Navdeep Suri, (Observer Research Foundation, 2021), https://www.orfonline.org/expert-speak/sovereignty-datafied-world-framework-indian-diplomacy/#_edn19.

38. Dustin Lewis, "An Enduring Impasse on Autonomous Weapons," *Just Security*, September 28, 2020,

39. https://www.justsecurity.org/72610/an-enduring-impasse-on-autonomous-weapons/.

40. K. Francis Park and Zur Shapira, "Risk and Uncertainty," The Palgrave Encyclopedia of Strategic Management, (Springer, 2017).

41. Stefan Kuhlmann, Peter Stegmaier and Kornelia Konrad, "The Tentative Governance of Emerging Science and Technology: A Conceptual Introduction," *Research Policy* 48, no. 5 (2019): 1091-1097.

Responsible Behaviour, Accountability and Consequences in the Chrome Age

James A. Lewis

A general lack of governance plagues the internet, but there are widely differing views on how to remedy this. There is deep disagreement among powerful states on how to proceed, particularly with regards to cybersecurity. This disagreement complicates the discussion of cyber norms. One complication was semantic, but reflected a deeper confusion. Norms are not implemented, they are observed. For a time, there was a discussion of how to "implement" norms, often accompanied by proposals for various new (and always redundant) norms. This discussion was not so much unhelpful as irrelevant. The issue is not implementation or proposing additional norms or even to clarify existing ones. The issue is to ensure that agreed norms are observed. Norms by themselves are insufficient. Changing this requires states to decide what to do when another state chooses to ignore what has been agreed to at the UN.

The most important norms for state behaviour in cyberspace are found in the UN Charter—respect for sovereignty, non-interference in internal affairs, refraining from the use of force or the threat to use force against another state. These foundational norms are frequently ignored. In response to increased concern about cyber threats to peace and stability, states developed additional norms for cybersecurity, and reached consensus on eleven specific norms developed by the 2015 UN's Group of Government Experts (GGE that define responsible state behaviour in cyberspace). While there are gaps in the 2015 norms, these are adequate to guide state action. These norms have since been endorsed by all UN member states at the conclusion of the Open-Ended Working Group (OEWG) in May 2021 (and the 2021 GGE, which concluded a month later, added useful clarifications). The OEWG agreement made these cyber norms politically binding.

Not all nations feel impelled, however, to observe them. A failure to observe norms is not addressed by creating additional norms, The failure reflects a national decision that norms can be ignored with impunity. It also reflects the absence of any consequence or penalty by the international community for a decision to ignore norms. Creating accountability, which may require the imposition of consequences, is the key to making agreed norms meaningful.

The need for penalties against violator nations points to the central role of States in making norms meaningful. States are uniquely placed to remedy the failure of another state to observe agreed norms. State efforts should be reinforced by civil society and the private sector, but experience shows that the countries that do not observe agreed norms are unlikely to be swayed by the actions of civil society or the private sector alone. State action is required. Close cooperation between civil society, the private sector, and governments is essential. This is not the case in authoritarian regimes, and while a failure to involve civil society and the private sector may ultimately weaken authoritarian states, the immediate responsibility for creating and enforcing codes and norms falls on states, who must develop the diplomatic, military and intelligence strategies to increase adherence to what has been agreed. Very little work has been done to create such strategies

This reflects long-established state practice on the use of force. Force, or the threat to use force, is the ultimate source of accountability against another state when persuasion has failed. Sovereigns have rights to use force that private actors do not. The centrality of states also reflects differing capabilities for the use of force and diplomacy, as states have institutions dedicated to the use of force and coercion.

Accountability

The GGE norms endorsed by the General Assembly were a first step towards building cybersecurity in a "rules-based" international environment, but agreement on norms is by itself not enough. Sovereignty is routinely violated through cyber intrusions (most recently demonstrated by China's "Hafnium" intrusion)[1]; coercion is routinely employed as part of a growing pattern of activity to interfere in the internal affairs of other nations (notably by Russia's frequent interference in national elections in Europe and the US). What we can conclude from these disturbing trends is that agreements that lack commitments from powerful states or that fly in the face of state practice will prove insufficient to improve the situation in cyberspace. For norms to have effect, there must be consequences for nations that choose not to observe them. States that are not willing to observe norms are likely to change their behaviour only if they assess the consequences of misbehaviour as unacceptable. The intent for the imposition of consequences is to increase the observation of norms and strengthen accountability.

Norms for responsible state behaviour in cyberspace will form part of the larger structure of digital governance, but only if they are given meaning through the development of consequences. Consequences mean the range of internationally lawful responses available to states that are the victim or target of malicious actions that run counter to agreed norms. While there has been some progress towards the imposition of consequences, this has so far been an ad hoc and disjointed process. Nor are sporadic responses sufficient to persuade hostile states that observation of norms is in their best interest.

Consequences fall into three general categories—diplomatic, economic, and the use of force. The task for diplomacy is to identify specific examples of actions within these categories, communicate them to shape opponent perceptions of risk, and ensure greater global acceptance of retaliation. This task is the next step for improving the observation of norms and building collective action.

This approach must be public. Simply announcing the development of a menu of consequences will not produce the needed effect. The reasons for a public discussion of consequences are to provide a basis for collective action among likeminded states and to socialise with the international community the idea of imposing legitimate, justifiable consequences for a failure to observe norms. The development and potential imposition of consequences must be part of a larger diplomatic strategy by likeminded governments for how international relations will change given the dynamics of the risk of new powers and the weakness (perhaps temporary) of the old.

Accountability, however, has always been a problem for the international system. The UN Security Council, designed to provide accountability, rarely can agree on consequences or penalties for a failure to observe norms of international behaviour. With rare exceptions, this has been true since Nikita Khrushchev thumped his shoe on the UN General Assembly's podium, if not before. The

backdrop and alternative to a rules-based order remains the stark truth of the Melian Dialogue, "that right, as the world goes, is only in question between equals in power, while the strong do what they can and the weak suffer what they must."[2] Any rules-based order must be flexible enough to accommodate the realities of international power. No powerful state will allow an international organisation to impose binding rules on activities it regards as its prerogative unless it chooses of its own accord to do so or is forced to accept this decision.

There has been slow progress towards reducing the Melian principle's influence in international relations and replacing it with the rule of law, but this is far from complete. In fact, the 1945 international system has for some time been becoming more fragile. In recent years, frequent appeals to a rules-based order built on international law and championed by the US and its allies probably reflect an awareness that this post-1945 rules-based order is in decline. The US's ability to promote accountability has weakened, but that does not translate into it being replaced by China, which, despite its wealth, lacks both an appealing alternative vision for the global order (unless one is entranced by Xi Jinping thought, and few are) or the diplomatic skills to attain preeminence. What China's rise does mean, however, is that our assumptions about the effectiveness of rules, the nature of stability, and environment for conflict, are increasingly open to question, creating a more difficult landscape for the observation of norms and for diplomacy.

This erosion of an American-led international order and China's desire to replace it shapes the global agreement on cybersecurity. Effective agreement on cybersecurity could help reverse this decline, as part of a larger effort to create new rules and institutions to govern the digital space. But any new international order will not be global. It will be based on agreement among a smaller, like-minded community of nations. Acceptance of the rule of law is a fundamental criterion for membership in this community. This means that a degree of bifurcation is unavoidable.

Accountability raises other issues. These include attribution and information-sharing, mechanisms for coordination, and agreement on proportionality. All will need to be developed by states to create a collective approach to accountability. The community of norms-observing states can create accountability by using the 2015 GGE norms as a framework for action when there is a failure to observe them. The goal is to reduce the number, scope, and risk of malicious cyber actions. To achieve this, they must be prepared to act collectively against those who do not chose to observe the agreed norms. This will require mechanisms for cooperation (probably informal) and common understandings of attribution, proportionality, and managing any risk from responsive action. The ultimate arbiters of responsible behaviour are other states, operating in the political process of diplomacy.

Attribution

Attribution of the source of a malicious action is essential. While attribution capabilities vary widely among countries, private sector actors, such as FireEye or Cloudstrike, now have the capability to attribute the source of a hostile cyber action. This progress in attribution is not recognised in the international community, and there is a lack of agreement on what level of attribution is required for cooperative action among states, but these issues can be remedied and should not prevent steps to increase accountability or impose consequences.

The discussion of cybersecurity has been hampered by confusion over what is meant by attribution. Some see it as technical and forensic; others see it as a legal construct. The chief obstacles are political. Attribution of the source of a malicious cyber action will remain a national decision and any construct for collective action must recognise this. A sovereign state has the right to decide who attacked it. States will not give up that right to an international attribution mechanism or some international arbiter for attribution, and such proposals are unlikely to ever find agreement. The International Atomic Energy Commission (IAEA) is not a precedent. It is based on a treaty that reflects political agreement among states that the development of nuclear weapons is unacceptable. There is no such agreement for cyber actions. The IAEA has technical standards for attribution based on the physical attributes of nuclear activity. Such standards do not exist for cyberspace. Attribution remains the prerogative of states and taking collective action will depend on one state persuading others of the validity of its case.

There is still a general tendency in cybersecurity to overvalue technical aspects of the discussion. We are not talking about technical attribution when it comes to collective action to impose consequences on states that fail to observe norms. Nor is this the kind of attribution required by a court of law. There are no judges in cyberspace, impartial or otherwise, and the evidentiary standards required in court are profoundly inappropriate for relations among states. The object is political attribution, a decision by a state that it is persuaded as to the identity of an attacker. There have been recent examples of political attribution, including European Union (EU) sanction on Russia for hacking the German Bundestag, and the identification by NATO, the EU, Australia, Britain, Canada, Japan, New Zealand, and the US that China was responsible for an enormous hack on Microsoft Exchange[3].

Proportionality

These recent actions are valuable first steps, but they raise questions of whether the response was effective. The ultimate measure of effectiveness is an observable change in hostile state behaviour (and this will not occur in a single instance, as if cybersecurity was a ping-pong match). Proportionality is also a requirement. States that observe international law will seek proportionality in any response. Defining a proportional response will also be essential for creating the agreement among states for collective action, but proportionality must be approached strategically. A tit-for-tat exchange is unlikely to be effective, given how long the US and others have allowed Russia and China to act without penalty.

Determining proportionality for a response to a malicious cyber action that is both effective and lawful remains difficult. Malicious action will be coercive, a violation of sovereignty, harmful, but will most likely not count as an armed attack or use of force. At most, it may involve the threat to use force. This makes simple equivalence in response inadequate. What is the proportional response to the hacking of an election by a country where elections are a sham? Retaliatory hacking of an election whose outcome is known well before any voting is pointless. An initial conclusion is that proportionality cannot be limited to cyber responses. A second is that our current understanding of proportionality may be inadequate and equivalence in response may need to be determined by the overall and cumulative pattern of malicious activity rather than a specific, individual event.

While it is possible to draw precedents from kinetic action to determine proportionality, this is insufficient. There is tacit agreement among active cyber powers to avoid cyber actions that cause physical destruction, damage, or casualties, and therefore do not qualify as the use of force, and while there is growing consideration of the parameters of "logical" destruction (such as erasure or disruption of data) there is not yet consensus on this. Kinetic precedents are inadequate for determining the proportional response to malicious cyber actions, particularly if they involve economic espionage or influence operations. If cyber actions destroy critical infrastructures, there is ample precent for response. If cyber action destroys a telecommunications manufacturer through the theft of intellectual property, the precedents are much less clear.

Proportionality and attribution are essential ingredients for a politically acceptable response to malicious cyber action—acceptable to domestic publics, to allies and to the global community. A proportional response will be seen as just, but proportionality is complicated because responses cannot be sporadic, reactive actions if they are to affect opponent behaviour. Promoting the observation of norms will require a strategy of sustained engagement for the imposition of consequences.

One dilemma created by decades of failure to respond to malicious cyber acts is that those who fail to observe norms do not believe they face any risk and dismiss the non-forceful actions preferred by democracies. Censure, indictments, and sanctions do not appear to have any effect. This stems from more than cyber operations. The failure to respond effectively in Ukraine, Syria, and the South China Sea created a perception that norms of international behaviour can be ignored. The immediate effect of this is that merely threatening to impose consequences will, at least at first, not be seen as credible.

Escalation

Responses to malicious cyber acts have been hampered by unreasonable concerns that any counter action will lead to an escalation of conflict. This has undercut the observation of norms. We now have sufficient experience with cyber conflict to know escalation is much less likely than was originally expected. The risk of escalation from cyber actions is exaggerated. The states with the most advanced cyber offensive capabilities treat cyber as a strategic asset and control it tightly. A fondness for hypothetic scenarios, a legacy of the Cold War nuclear fears, contributes to miscalculation. There has never been a cyber incident that resulted in the escalation of conflict.

The reasons for this are suggestive for the nature of international conflict in the first decades of the twenty-first century. Escalation is unlikely in part because nuclear powers, even in South Asia, have been exceptionally reluctant to use nuclear weapons, despite possessing them for decades. Despite occasional nuclear saber-rattling by authoritarian, states are cautious and unwilling to risk nuclear war[4]. The existential nature of nuclear war makes it difficult to coolly assess risk. Even below the nuclear threshold, however, conventional war between major powers is too damaging and too expensive to be sustained for long periods. We can test this by looking at conventional conflicts over the last decade that have tended to be short and limited in scope (the US's costly misadventures in the Middle East do not qualify, as these were against poorly armed, irregular opponents).

Precision guided munitions and unmanned aerial vehicles, guided by greatly improved reconnaissance capabilities, have changed warfare, and make it more costly. The mix of precision-guided munitions unmanned aerial vehicles and cyber actions allow for states to create strategic effect without using nuclear weapons. Developments in hypersonic delivery vehicles and artificial intelligence-enhanced weapons will further expand this capacity for non-nuclear strategic effect. National strategies have not fully accommodated this technology-driven change in how wars will be fought. The cost of nuclear and conventional conflict is one reason why cyber operations are so attractive; the risks of retaliation for cyber actions are disproportionately small compared to the benefits, but it is also one reason why nations are careful to avoid escalation risks.

Manipulating the fear of escalation is a standard tactic for some authoritarian states to inhibit any response to malicious cyber action. Even if common understandings on attribution and proportionality are developed, there will need to be a greater acceptance of risk by democratic states if norms are to be made effective. Risk intolerance in enforcement actions may be the greatest impediment to effective norms.

Fears about retaliation have ceded the initiative to less risk-averse, authoritarian states and have handicapped the development of effective responses. Miscalculation of the risk of escalation reflects the overly powerful influence of strategic concepts inherited from the twentieth century and the threat of global nuclear war, It damages security to apply the template of twentieth century strategy to the problems of the twenty-first century without recognising the constraints that nuclear weapons have created and the changes and new technologies have brought for interstate conflict among nuclear powers.

The concept of deterrence, which dominates much of US thinking, may undercuts the observation of norms. How does one deter a failure to observe norms? Deterrence involves certain symbolic actions—amassing missiles, making declarations of redlines—intended to shape opponent calculations of risk, but the immense difference in risk between cyber action and nuclear action means that deterrence symbolism lacks the heft and credibility needed to affect cyber opponents. If it were possible to credibly associate a nuclear response to a cyber action, deterrence might work, but a nuclear response to any cyber action undertaken would be unacceptably disproportionate and ensure a much greater risk of retaliation.

Consequences and Digital Diplomacy

Cyber operations have become a permanent and increasingly important element of international conflict. How countries use cyber operations is determined by their larger interests, by their existing strategies, experience and institutions, and by their tolerance for risk, but many have experiments and more have developed the capabilities need for offensive operations. This means that developing effective rules, or at least shared understandings among major cyber powers on use and norms has become a vital diplomatic task.

Changing opponent perceptions of the risks of taking cyber actions contrary to agreed norms offers the best avenue to increased stability, or at least decreased levels of malicious actions. For norms to have effect, there must be consequences for nations that choose not to observe them. Restoring credibility with the intent of giving norms effect will require the development

of a menu of consequences that are politically acceptable to a global audience and effective in changing behaviour. Credible threats of punishment for a failure to observe norms is the only tool available since authoritarian states are not going to disappear or of their own accord decide to end their cyber operations.

Ultimately, the normative framework agreed in 2015 and 2021 can reinforce the structures of international law and universal norms to reduce cyber conflict. The global agreement on a framework of responsible state behaviour was a significant step forward for building international cybersecurity in a rules-based environment, but experience has shown that agreement on norms is not enough. The digital environment has become a primary location for interstate conflict. Reducing the level of conflict requires not only norms but consequences for the failure to observe them. This will not be an easy task or without risk, but it is essential as we unavoidably move into a world of increased conflict between authoritarian and democratic states.

Endnotes

1. "What is the Hafnium Microsoft hack and why has the UK linked it to China?," *The Guardian*, July 19, 2021, https://www.theguardian.com/world/2021/jul/19/what-is-the-hafnium-microsoft-hack-and-why-has-the-uk-linked-it-to-china

2. Thucydides, "CHAPTER XVII. Sixteenth Year of the War - The Melian Conference - Fate of Melos," in *History of the Peloponnesian War*, (South Hadley, Mount Holyoke College) https://www.mtholyoke.edu/acad/intrel/melian.htm.

3. Laurens Cerulus, "EU sanctions Russian hackers for 2015 Bundestag breach," Politico, October 22, 2020, https://www.politico.eu/article/eu-sanctions-russias-fancy-bear-hackers-for-2015-bundestag-breach/; Eric Tucker, "Microsoft Exchange hack caused by China, US and allies say," AP News, July 19, 2021 https://apnews.com/article/microsoft-exchange-hack-biden-china-d533f5361cbc3374fdea58d3fb059f35.

4. Nina Tannenwald, "23 Years of Nonuse Does the Nuclear Taboo Constrain India and Pakistan?," *Stimson Center*, February 22, 2021, https://www.stimson.org/2021/23-years-of-nonuse/#.

Narratives

US vs Big Tech: Ten Trends Pointing to a Fundamental Relationship Reboot

Nikhila Natarajan

Four antitrust cases against Google[1] and one more getting cooked, two against Facebook, a 15-month investigation by the US Congress, a scathing 450-page report[2], an executive order on reining in online market power, Apple and Amazon under scrutiny, and plenty of 'techlash' bills gathering momentum at the state level[3]. Groundbreaking action is advancing in the US—in the courts, in Congress *and* via the White House—against online market power. Hostility seems to be intensifying, and the digital surface area covered by the potential fallout has never been higher.

Then there is the political calculus. Democrats control the House, but they need Republican support in the Senate for legislation to pass. In the current 50-50 Senate,[4] every senator has leverage, every contentious bill—Big Tech *or* otherwise—is one vote away from failure.

Despite political gridlock, it is increasingly clear that Republicans and Democrats have found common ground: Big Tech has gotten too big, the harms to children are unacceptable, our psychological levers are getting twisted, our vulnerabilities exploited, businesses that undermine public health, democracy and privacy are a clear and present danger. US lawmakers are drawing straight lines from the algorithmic features of surveillance capitalism to the Cambridge Analytica scandal during the 2016 presidential election[5], to the storming of the US Capitol on 6 January 2021, to COVID-19 vaccine misinformation and, therefore, life and death. By July 2021, nearly all COVID-19 deaths in the US—a country that has an excess supply of the shots—were among the unvaccinated. In a 22-page advisory, US President Joe Biden's surgeon general warned that the pace of COVID-19 vaccinations has slowed throughout the country, riding on vaccine opposition fueled by false claims on social media about the safety of shots. The US death toll crossed 600,000 this summer and by September, the surge in cases and hospitalisations in the least vaccinated states overwhelmed medical workers angry and exhausted by the full circle repeat of the 2020 nightmare scenarios. It was smack in the middle of the pandemic's catastrophic march in 2020 when a steady string of Congress hearings on Big Tech went quickly from novel to routine. Time and again, Google's Sundar Pichai, Facebook's Mark Zuckerberg, Amazon's Jeff Bezos, Twitter's Jack Dorsey, and Apple's Tim Cook appeared in video windows, for hours on end. Each time they reappeared on our screens, their talking points were refined by the previous encounter, their arguments tactically altered for the looming collision course with lawmakers and the courts, with one eye constantly on the markets.

It is a severe contest. What does the fall of Big Tech, from Capitol Hill darlings to anti-trust targets, say about Washington DC's changing relationship with platforms? Ten headline themes mark Big Tech's souring relationship with Washington DC:

Wheels Turn Slowly

The US's antitrust laws first came into being more than 100 years ago. It was a different, pre-digital era in American history, where the problem of 'big' was about the "imbalance of power between

the rich and those of more modest means" [6]. Matt Stoller, director of research at the American Economic Liberties Project, notes how slowly the wheels turn on antitrust abuses. Microsoft is a case in point. In his book *Goliath*[7] , Stoller writes about how, taking the cue from the companies' rivals, the US Federal Trade Commission (FTC) first began investigating Microsoft over antitrust abuses in 1991 and closed the investigation in 1993, after which the Justice Department began its own probe. It took until 2000 for the court to rule for the government and lay out a plan to break up Microsoft. Gates appealed the decision, an ultra-conservative court overturned the breakup order, and in 2001, the George Bush administration dumped the case. In between all this, there was a moment when Gates told *Businessweek,* "The worst that could come out of this is that I could fall down on the steps of the FTC, hit my head and kill myself." Noah Philips, one of the five current commissioners at the FTC, gives us a sense of the current mood within the organisation: "We have a lot of taking down the old and not a lot of guidance as to what the new will be. And that is the ultimate question, right? what is the policy going to be and what is the basis…?"[8]

New Linchpins, New Vocabulary

On 9 July 2021, the White House made public unprecedented aggression on Big Tech and antitrust[9]. The vocabulary shifted; the mood thereon rooted in activist ethos. On that day, the Joe Biden administration confirmed that it is ready to play hardball. At its core, Biden's Big Tech playbook seems perched on three pillars—killer acquisitions; surveillance and the accumulation of data; and unfair methods of competition on internet marketplaces.

These terms are no longer the preserve of academic research papers. We see them mainstreamed in the text of executive orders, investigations, and policy priorities. The people driving these policies and their patois signal an unmistakable shift. At the tip of the spear are Lina Khan, who at 32 is the youngest chair in the history of the FTC, and Timothy Wu, who serves as Special Assistant to the President for Technology and Competition Policy. A freeze frame that captures the reboot is Biden passing one of his shiny black Cross Century II Rollerball pens to Khan[10]. The swishy presidential signature that afternoon promoted competition throughout the US economy. Biden seems to have embraced the Khan-Wu tonality. "No more tolerance for abusive actions by monopolies. No more bad mergers that lead to mass layoffs, higher prices, fewer options for workers and consumers alike," he said at the ceremony before signing the order. Biden's initial argument, via the executive order, is that dominant internet platforms are destroying new players and, as a result, can extract monopoly profits and gather and exploit intimate personal information. These themes derive from and are hardcoded into the US Congress's 450-pager scholarly and legal pushback. Broadly, the Khan-Wu view is that federal agencies must lean into their jobs of investigating, reporting, and creating rules of the road for big business. The new executive order goes there. This got hardcoded just a bit more in the Apple versus Epic litigation September, where without taking the game away completely, a California judge etched a serious fault line in the famed payment fortress that Steve Jobs and Tim Cook built around app developers[11]. The matter will go in appeal, perhaps from both sides. Epic's litigation against Google has reached trial stage too.

"Buy or Bury": The Antitrust Attack Against Facebook

On 28 June 2021, federal judge James E. Boasberg issued a stunning rebuke of the government's efforts to break up Facebook over alleged antitrust violations. After getting roasted in round one, federal regulators have come back with all guns blazing in round two of their antitrust attack on Facebook[12]. Led by Khan, the US government now alleges that the social network giant pursued a "buy or bury" strategy against rivals. Khan seems to be seeking remedies that include "divestiture of assets, divestiture or reconstruction of businesses (including, but not limited to, Instagram and/or WhatsApp)". It is instructive here to track the basis on which the judge threw out the earlier complaint (but not the case), calling it "legally insufficient" to prove that Facebook was a monopoly. In two opinions of more than 50 pages each, the justice said that in one case, the FTC failed to define the market that Facebook operates in and another case that a coalition of 48 state attorneys general sat on the sidelines for too long[13]. As Facebook headquarters erupted in celebration, the FTC went back to work, taking the cues offered in the rebuke and sharpening its attack. In the new filing, the FTC laid out a detailed analysis to substantiate its monopoly power claim against the social networking giant. FTC voted 3-2 to file the amended complaint, with two Republicans voting against it[14].

Free Speech and "Poison"

The train wreck is not limited to the storming of the US Capitol on 6 January 2021 or the 2016 Cambridge Analytica scandal, and extends beyond privacy and competition. As an explosive anti-COVID-19 vaccination campaign took wing across the country, US Surgeon General Vivek Murthy urged social platforms to slow the spread of the 'poison'[15]. Murthy tore into dominant tech companies for enabling COVID-19 misinformation and called on them to redesign their recommendation algorithms and construct built-in "frictions".

"Modern technology companies have enabled misinformation to poison our information environment with little accountability to their users," Murthy declared at a White House briefing this summer. "We are asking them to step up, we can't wait longer for them to take aggressive action"[16].

The White House called on Facebook to remove 12 accounts that may be responsible for as much as 65 percent of COVID-19 disinformation. "They're killing people," Biden said about Facebook, before dialing it down a day later[17]. "Misinformation poses an imminent and insidious threat to our nation's health," Murthy said. "We must confront misinformation as a nation. Lives are depending on it"[18]. It is the first time the perils of social media recommendation algorithms were featured in a Surgeon General's briefing. It will not be the last.

Section 230 of the 1996 Communications Decency Act—which is part of a larger telecommunications law—provides platforms protections from liability for the content users post. Twenty-six words in Section 230—"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"—make social media, the way we know it, possible[19]. The question on where to draw the line continues to rage on. But who (all) should be drawing that line?

Data Extraction

The pushback against secret data extraction seems real. The purported illegitimacy, the process, the subterfuge, and its outsize value to companies, all seem to be on the table. "The fundamental act of unilateral extraction from our lives is deeply illegitimate. And of course, without that, the whole rest of the edifice crumbles. When we talk about objectives in the larger chessboard, we need to be able to confront extraction head on. And we need to have laws that make this kind of secret extraction illegal. Simply criminalize it," Shoshana Zuboff, author of *Surveillance Capitalism*, said in a 2020 interview[20]. "America needs a federal data protection law right now," declared Marietje Schaake, international policy director at Stanford University's Cyber Policy Center[21].

Ex-Facebook staffer Dipayan Ghosh points us to the "minimal" marginal technological cost of collecting data. The asymmetry of knowledge, he argues, left consumers well behind the industry. "The most striking feature of this terrible new circumstance is that the algorithmic machine has matured organically. It is not the result of a concerted business plan but rather an experimentally and empirically evolved animal trained to identify opportunities for economic arbitrage in the novel industry of manipulated communication," Ghosh writes in *Terms of Disservice*[22].

A patchwork of privacy legislation has begun to emerge in the US, but comprehensive privacy laws are yet to be passed at the federal level. States are leading. California's Consumer Privacy Act is out there, other states are following in that general direction. In an opinion piece headlined 'How Capitalism Betrayed Privacy', Wu writes, "Mass privacy is the freedom to act without being watched and thus, in a sense, to be who we really are - not who we want others to think we are. At stake then, is something akin to the soul" [23].

The privacy issue is "far from being resolved", write Michael Kearns and Aaron Roth in *The Ethical Algorithm*. "…But at least it is in a state where we can speak rigorously about the kinds of problems that can and cannot be solved by current technology" [24].

Interoperability

What does genuine interoperability look like on social platforms? What should it look like? Lawmakers are pushing for answers. Big Tech does not want to make it terribly easy. "Are you willing to give me the right to take my data on Facebook and move it to another social media platform?" Zuckerberg was asked during a 2018 testimony in US Congress. "Senator, you can already do that," he replied. "We have a 'Download Your Information' tool where you can go and get a file of all the content there and then you can do whatever you want with it" [25].

Those who have ever tried this know better. What you get is a text file loaded with your friends' names and their joining dates. "Breaking up social media without ensuring interoperability is perhaps the worst of all possible approaches," writes Sinan Aral in *The Hype Machine*. One of the many challenges he mentions is what he calls the "transparency paradox"—the tensions between being more open and cranking up security at the same time, between free speech and data protection, privacy integration and election integrity.

The Media

In the US, there seems to be growing cognisance of how disproportionate revenue sharing between Big Tech and media publishers is hollowing out the local news industry. In 2019, just 16 percent of the US $134 billion that advertisers spent through Google went to more than two million non-Google properties[26], which includes news publishers that also sell their advertising space through Google's exchange and buying tools. Approximately 25 percent of all newspapers in the US have merged or gone out of business in the past 15 years. Between 2018 and 2020, 300 newspapers went bust and 6,000 journalists lost their jobs[27]. US lawmakers are getting more involved. The Journalism Competition and Preservation Act has been reintroduced for the third time since 2018[28]. If it makes headway, the bill will offer a pathway for news publishers to negotiate as a group with "dominant online platforms" for a larger share of online advertising revenue. News publishers are supporting this effort, along with the Local Journalism Sustainability Act[29], which proposes annual US$250 tax credits to encourage Americans to subscribe to local news or donate to local nonprofit news organisations. That bill—which has caps of different dollar amounts for news subscribers, publishers, and small businesses that advertise in local newspapers—is gaining traction on Capitol Hill and has picked up 70 cosponsors. The rising wave of legislation specific to the news industry adds to the growing pile of so-called "techlash" bills in statehouses, where there is a growing bipartisan effort to install new guardrails against Big Tech platforms around matters related to antitrust, consumer privacy, and public interest. The politics are tough, positions are hardening, and the fight is on to frame a new narrative.

Harms to Children

Child-oriented apps and manipulative influencer marketing that profit from a vulnerable section of the population are in the firing line. Lawmakers are blaming Big Tech for hijacking children's minds and for steering them toward Big Tech goals. If kids are scrolling at 2 am instead of sleeping, that conversation is now spilling over into Congress hearings; hearings like "Protecting Kids Online: Internet Privacy and Manipulative Marketing"[30] are now par for course. Attorneys general, civil society, and lawmakers are dashing off letters to Big Tech, calling for them to slam the brakes. Facebook's plans to launch Instagram for children under the age of 13 set alarm bells ringing. "It appears that Facebook is not responding to a need, but instead creating one, as this platform appeals primarily to children who otherwise do not or would not have an Instagram account," said a letter, signed by the attorneys general of 40 states, the District of Columbia and three US territories[31]. In a bombshell report on 14 September 2021, *Wall Street Journal* reported that Facebook researchers have repeatedly found that Instagram is toxic for teen girls and played it down in public. "Aspects of Instagram exacerbate each other to create a perfect storm," according to Facebook researchers quoted in the report[32].

"It's not okay to have to hide your own life, it is not normal…" Zuboff writes in *The Age of Surveillance Capitalism*[33]. "When I speak to my children or an audience of young people, I try to alert them to the historically contingent nature of the "thing that has us" by calling attention to ordinary values and expectations before surveillance capitalism began its campaign of psychic numbing."

Fundamentally, this conversation—and the larger one—is about the misalignment between the business model and what is best for people. What are the harms from the current digital advertising model and what are the alternatives? Who knows? Who decides? Who decides who decides? *The Social Dilemma*, among Netflix's most popular documentaries, is a superhit cautionary tale about social media design and the resulting addiction, and has used the internet's phenomenal reach to touch upon the dire consequences of surveillance, capitalism, addiction, and polarisation. The story is told through the life and times of a family, with teens and self-esteem issues at its core. The cast of *The Social Dilemma* is now a regular fix in Senate hearings in Washington. Lawmakers refer to the film often, their lexicon reflects new layers of knowledge gathered over the last year alone. Netflix just dropped the film on YouTube, now playing without a paywall. Big Tech's big defectors have gone from being the industry's champions to its sharpest critics. Roger McNamee, Tristan Harris...the A-list keeps growing.

Meanwhile in Silicon Valley

Incredible earning, bumper profits and the new product buildout is relentless. Roughly 30 percent of the global population uses at least one Facebook app every single day. Facebook, Alphabet (Google's parent company) and Microsoft reported strong Q2 numbers. In the case of Facebook and Google, both reported topline increases of more than 50 percent—Google at 62 percent [34] and Facebook with 56 percent[35]. YouTube, with two billion monthly average users, is logging one billion hours of video watched every day. Businesses of all sizes are merrily paying Facebook and/or Google for better ad placement, return on ad spend, and measurement of ad effectiveness. Against the backdrop of all the hand wringing around Big Tech's digital advertising model, which provides the bulk of the industry's profits, Zuckerberg is already outlining a future for his company that is *not* based on advertising. In a remarkable shift, he is talking about building the "metaverse", a Silicon Valley buzzword pointing to the next decade of innovation in the digital economy. In Zuckerberg speak, the metaverse is "a virtual environment where you can be present with people in digital spaces." It is "an embodied Internet that you're inside of rather than just looking at. We believe that this is going to be the successor to the mobile Internet" [36]. While lawmakers are trying to build guardrails against harms from the current breed of computer mediated communication channels, Facebook is already drawing up visions of a whole new world of virtual real estate. Industry chatter suggests the metaverse is being designed to envelop us in a more immersive grip than we have experienced before. Big Tech's fiercest critics are urging consumers to see that the convenience of internet platforms increases vulnerability, and it is a terrible thing because it undermines the most meaningful personal relationships.

Break Them Up

Riding on damning notes from internal emails—"Instagram is eating our lunch", "One thing about startups though is that you can often acquire them", "How do we deal with the threat of proliferating verticals"[37]—US lawmakers have been stitching up an elaborate case against tech giants who run platforms and, in parallel, juice the resulting market intelligence to operate as competitors on the same platform[38]. Ten rounds of such "exhibits" formed the bedrock of a string of Big Tech hearings in the US Congress, which kicked off in the summer of 2020. Over a year, the trajectory of the attack has come into sharper focus. The clearest signals arrived a year

later. On 24 June 2021, the US House Judiciary Committee greenlighted the final piece of a big box package, the Ending Platform Monopolies Act[39]. The name makes the intent clear. At a high level, the package aims to slam the brakes on Big Tech's ability to own a dominant platform in parallel with another line of business if those two create a conflict of interest. Without naming Facebook, Amazon, Google or Apple, the package refers to new guardrails for a certain category of online platforms—50 million or more monthly active users; annual sales or market value of over US$600 billion; and a role as a "critical trading partner." Not everybody is convinced. Lawmakers from Silicon Valley liken such action to a "grenade" that will simply blow up the tech economy[40]. What the bills in the larger package attempt to do:

Ending Platform Monopolies Act: prevents platforms with a market cap of US$600 billion and over 50 million monthly US users from operating businesses that compete with users on the platform.

American Choice and Innovation Online Act: prohibits Big Tech from favouring its own products and services over those of its competitors (for instance, Google pushing Yelp reviews further down on search results). Also, Big Tech cannot use data on its platform to create competing products, as in the Amazon versus Diapers.com case [41].

Platform Competition and Opportunity Act: the burden of proof will be on Big Tech to prove that future acquisitions are not unlawful. At present, the US government must do so.

Augmenting Compatibility and Competition by Enabling Service Switching (ACCESS) Act: pushes Big Tech to make data portability and interoperability convenient for users so they can switch between platforms.

Merger Filing Fee Modernization Act: when large mergers are proposed, the FTC and Department of Justice get paid for poring over the details. The bill raises that fee to better fund departments that do antitrust work.

Conclusion

Setting the pace in all this is no longer an American preserve. The European Union (EU), for example, has long outpaced the US on privacy regulation. In early September, Ireland's privacy watchdog fined WhatsApp a record US$267 million[42] for breaching EU data protection rules on sharing people's data with other Facebook companies. This is the second largest fine in the EU under the General Data Protection Regulation, behind Luxembourg's US$887 million fine to Amazon, also for data protection violations[43].

Scholars like Zuboff, Ghosh and Aral advocate strongly for the role of the US government in protecting privacy and taking the fight to monopolies; they believe that done right, this does not crush innovation but instead helps create a new digital social contract that puts the individual at its core. The US Congress testimonies of Zuckerberg, Dorsey, Pichai and others make clear that charting rules of the road for a common technological future is going to be incredibly complex, technically deep and all very new. Like Zuboff reminds us, this is the first century in which we have experienced digital and "unprecedented harms means that we need new kinds of solutions"[44].

The pressure is growing because of action elsewhere in the world, including the three Quad partners (India, Australia, and Japan). An illustrative example is in media regulation[45]. On 25 February, 2021, Australia passed the News Media Bargaining Code, a world-first law that forces Internet giants Google and Facebook to pay for news content shared on their platforms. The country instantly became a testing ground for digital platform micro regulation. Acting independently, at least thus far, the governments of India and Japan are, like Australia, contemplating policy interventions at the intersection of news media, social media platforms, and the ethics of digital media intermediaries. Partner nations might follow suit with their own version. "Now the best we can do is to steer this momentum in the right, more democratic direction," Schaake said in the aftermath of the 6 January attack on the US Capitol. In a word, the US's point seems to coalesce with the line Khan and Wu took as academics.

"In my dreams, billions of consumers will rebel," writes Roger McNamee in *Zucked: Waking up to the Facebook Catastrophe*[46]. "Nowhere is it clearer that we need presidential leadership to take actions and change laws and lead investigations than in Big Tech," former US presidential candidate Amy Klobuchar writes in her 2021 book *Antitrust: Taking on Monopoly Power From the Gilded Age to the Digital Age*. "This is finally real."[47]

Endnotes

1. Lauren Feiner, "States Bring a New Antitrust Suit against Google over Its Mobile App Store," CNBC, July 7, 2021, https://www.cnbc.com/2021/07/07/states-bring-new-antitrust-suit-against-google-over-google-play.html.

2. US Congress House Committee on the Judiciary, Documents from the Hearing on "Online Platforms and Market Power: Examining the Dominance of Amazon, Apple, Facebook and Google, Washington D.C, 2020, https://judiciary.house.gov/online-platforms-and-market-power/.

3. Ohio Attorney General Dave Yost, https://www.ohioattorneygeneral.gov/Media/News-Releases/June-2021/AG-Yost-Files-Landmark-Lawsuit-to-Declare-Google-a

4. Jordain Carney, "Manchin Cements Key-Vote Status in 50-50 Senate," *The Hill*, March 11, 2021, https://thehill.com/homenews/senate/542664-manchin-cements-key-vote-status-in-50-50-senate.

5. C-SPAN, "Cambridge Analytica and Data Privacy," Online Video, 98:57 min, May 16, 2018, https://www.c-span.org/video/?445621-1/cambridge-analytica-whistleblower-christopher-wylie-testifies-data-privacy

6. Amy Klobuchar, *Antitrust: Taking on Monopoly Power from the Gilded Age to the Digital Age* (Knopf, 2021).

7. Matt Stoller, *Goliath: The 100-Year War Between Monopoly Power and Democracy* (Simon & Schuster, 2019).

8. Hudson Institute, "Perspectives on the Current Federal Trade Commission," YouTube video, 49:32 min, August 30, 2021, https://www.youtube.com/watch?v=n1OjLC8GZqU

9. The White House, Executive Order on Promoting Competition in the American Economy, Washington D.C, 2021, https://www.whitehouse.gov/briefing-room/presidential-actions/2021/07/09/executive-order-on-promoting-competition-in-the-american-economy/.

10. The White House, "President Biden Delivers Remarks and Signs an Executive Order", YouTube video, 18:26 min, July 9, 2021, https://www.youtube.com/watch?v=vYzIwwtKXvo.

11. United States District Court, Northern District of California, Rule 52 Order After Trial, Epic Games Inc. vs Apple Inc, https://www.documentcloud.org/documents/21060696-epic-v-apple-ruling

12. Federal Trade Commission, Government of the United States of America, https://www.ftc.gov/news-events/press-releases/2021/08/ftc-alleges-facebook-resorted-illegal-buy-or-bury-scheme-crush.

13. Leah Nylen, "Federal Court Tosses Antitrust Suits against Facebook, in Huge Blow to D.C.'s Fight with Tech," *POLITICO*, https://www.politico.com/news/2021/06/28/federal-court-dismisses-ftcs-antitrust-case-against-facebook-496764.

14. Facebook has until 4 October 2021 to respond.

15. U.S. Department of Health & Human Services, Government of the United States of America, https://www.hhs.gov/about/news/2021/07/15/us-surgeon-general-issues-advisory-during-covid-19-vaccination-push-warning-american.html.

16. U.S. Department of Health & Human Services, Government of the United States of America, https://www.hhs.gov/about/news/2021/07/15/us-surgeon-general-issues-advisory-during-covid-19-vaccination-push-warning-american.html.

17. Lauren Egan, "'They're Killing People': Biden Blames Facebook, Other Social Media for Allowing Covid Misinformation," *NBC News*, July 16, 2021, https://www.nbcnews.com/politics/white-house/they-re-killing-people-biden-blames-facebook-other-social-media-n1274232.

18. U.S. Department of Health & Human Services, Government of the United States of America, https://www.hhs.gov/about/news/2021/07/15/us-surgeon-general-issues-advisory-during-covid-19-vaccination-push-warning-american.html.

19. "47 US Code § 230 - Protection for Private Blocking and Screening of Offensive Material," LII / Legal Information Institute, accessed September 7, 2021, https://www.law.cornell.edu/uscode/text/47/230.

20. "Secret Extraction of Behavioral Data Must Be Criminalized: Shoshana Zuboff," ORF America, https://orfamerica.org/recent-events/secret-extraction-of-behavioural-data-must-be-criminalized-shoshana-zuboff.

21. Nikhila Natarajan, "US Needs a Federal Data Protection Law Right Now: Marietje Schaake," ORF, January 29, 2021, https://www.orfonline.org/expert-speak/us-needs-federal-data-protection-law-right-now-marietje-schaake/.

22. Dipayan Ghosh, *Terms of Disservice: How Silicon Valley Is Destructive by Design*, Washington, D.C, Brookings Institution Press, 2020).

23. Tim Wu, "How Capitalism Betrayed Privacy," *The New York Times*, April 11, 2019, https://www.nytimes.com/2019/04/10/opinion/sunday/privacy-capitalism.html.

24. Michael Kearns and Aaron Roth, *The Ethical Algorithm: The Science of Socially Aware Algorithm Design* (New York: Oxford University Press, 2019),

25. Dylan Byers, "Senate Fails Its Zuckerberg Test," *CNNMoney*, April 10, 2018, https://money.cnn.com/2018/04/10/technology/senate-mark-zuckerberg-testimony/index.html.

26. US Congress House Committee on the Judiciary, Documents from the Hearing on "Online Platforms and Market Power: Examining the Dominance of Amazon, Apple, Facebook and Google - Exhibits, Washington D.C, 2020. https://judiciary.house.gov/uploadedfiles/0009.pdf.

27. Penelopy Abernathy and Zach Metzger, "News Deserts and Ghost Newspapers: Will Local News Survive?" UNC Hussman School of Journalism and Media, June 2020, https://www.usnewsdeserts.com/wp-content/uploads/2020/06/2020_News_Deserts_and_Ghost_Newspapers.pdf.

28. S. 673 - Journalism Competition and Preservation Act of 202", 117th Congress (2021-2022), https://www.congress.gov/bill/117th-congress/senate-bill/673.

29. H.R. 7640 - Local Journalism Sustainability Act, 116th Congress (2019-2020) https://www.congress.gov/bill/116th-congress/house-bill/7640?s=1&r=1.

30. US Senate Committee on Commerce, Science, & Transportation, Hearing on "Protecting Kids Online: Internet Privacy and Manipulative Marketing," May 18, 2021, https://www.commerce.senate.gov/2021/5/protecting-kids-online-internet-privacy-and-manipulative-marketing.

31. National Association of Attorneys General, USA, https://www.mass.gov/doc/naag-letter-to-facebook/download

32. Georgia Wells, Jeff Horwitz and Deepa Seetharaman, "Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show," *Wall Street Journal*, September 14, 2021, sec. Tech, https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-11631620739.

33. Shoshana Zuboff, T*he Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2019).

34. Jennifer Elias, "Google Advertising Revenue Rises 69% from Last Year," CNBC, July 27, 2021, https://www.cnbc.com/2021/07/27/alphabet-googl-earnings-q2-2021.html

35. Jordan Novet and Rodriguez Salvador, "Facebook Beats Earnings Expectations, but Warns of Significant Growth Slowdown," *CNBC*, July 28, 2021, https://www.cnbc.com/2021/07/28/facebook-fb-earnings-q2-2021.html.

36. Mark Zuckerberg, Facebook post, July 28, 2021, https://www.facebook.com/4/posts/10113801385499621

37. US Congress House Committee on the Judiciary, Documents from the Hearing on "Online Platforms and Market Power: Examining the Dominance of Amazon, Apple, Facebook and Google - Exhibits, Washington D.C, 2020. https://judiciary.house.gov/uploadedfiles/0009.pdf.

38. US Congress House Committee on the Judiciary, Documents from the Hearing on "Online Platforms and Market Power: Examining the Dominance of Amazon, Apple, Facebook and Google, (2020), https://judiciary.house.gov/online-platforms-and-market-power/.

39. H.R.3825 - Ending Platform Monopolies Act, 117th Congress (2021-2022), https://www.congress.gov/bill/117th-congress/house-bill/3825/text?r=34&s=1

40. "Big Tech Breakup on the Way? 5 Changes Congress Wants to See among Tech Giants," *Associated Press*, June 26, 2021, https://www.usatoday.com/story/tech/news/2021/06/26/how-congress-trying-rein-in-amazon-google-facebook-apple/5357023001/.

41. Timothy Lee, "Emails detail Amazon's plan to crush a startup rival with price cuts", *Ars Technica*, July 31, 2020, https://arstechnica.com/tech-policy/2020/07/emails-detail-amazons-plan-to-crush-a-startup-rival-with-price-cuts/

42. Celine Castronuovo, "WhatsApp Fined Record $267M over Breach of European Data Protection Rules," Text, *TheHill*, September 2, 2021, https://thehill.com/policy/technology/570623-whatsapp-fined-record-267m-over-breach-of-european-data-protection-rules.

43. Sam Shead, "Amazon Hit with $887 Million Fine by European Privacy Watchdog," *CNBC*, July 30, 2021, https://www.cnbc.com/2021/07/30/amazon-hit-with-fine-by-eu-privacy-watchdog-.html.

44. "Secret Extraction of Behavioral Data Must Be Criminalized."

45. Nikhila Natarajan, "The Quad, the Media, and Big Tech Platforms." ORF America, (2021), https://orfamerica.org/newresearch/the-quad-the-media-and-big-tech-platforms.

46. Roger McNamee, *Zucked: Waking Up to the Facebook Catastrophe* (New York: Penguin Press, 2019): 335.

47. Amy Klobuchar, *Antitrust: Taking on Monopoly Power from the Gilded Age to the Digital Age* (Knopf, 2021): 315.

It is Not Just the State, It is a Complicated Ecology

Paul Cadario

It is impossible to keep track of a complicated 'cyber-ecology' that changes every day.

With each new revelation about platform and device makers' enhancements, changes and upgrades, pundits and governments opine on their timeliness and merit. With each new investigation of privacy intrusions, users and suppliers demand reform, and politicians across the spectrum propose new legislation or demand closer regulation. And with each privacy or security breach, experts argue over whether it was the hardware or the software that was to blame or the user clicking on something they knew better than to be curious about. And every time a device or piece of software announces an update, a user has to say 'yes' and hope that this time it works.

The ecology is complicated: hardware (chips, devices big and small, servers that now serve "the cloud", and internet infrastructure), software (coding that makes applications and platforms work, and gives them their uniqueness and commercial appeal), the apps and platforms (relying on user trust for their popularity and profitability), and laws and regulations (that prescribe security and privacy requirements that providers must meet and most users take on trust as they accept user agreements and hope for the best).

Our modern world depends on hardware, some tiny and easily ignored, others visible and heavily marketed, working flawlessly. It relies on platforms, of which social media are the most widely used and discussed, and applications. Apps and platforms are all the product of complex, sophisticated, and sometimes flawed coding, hitherto largely by humans but increasingly enabled or performed by artificial intelligence (AI) and corrected promptly and universally when flaws are discovered.

Trust in All

Underpinning all of this is a trust factor—that the manufacturers have built physically secure widgets designed and sealed in devices that perform as intended and are secure from malware; that devices comply with standards that not just prevent them from overheating in your bag, pocket or office, but also ensure user privacy and network security; that network and infrastructure operators protect their customers from intrusions, snooping, and attack; and that governments and other regulators like the European Union (EU) or the International Telecommunication Union, adopt a Goldilocks 'just right' approach to regulation (neither 'too hot' nor 'too cold') to encourage innovation and competition while protecting users' data, communication and privacy.

The pandemic has brought this trust and reliability issue to the fore. Remote work has challenged organisations' IT staff to maintain connectivity and data security at home, and not just in the office, where the staff could stop around to diagnose, handhold, and service. Meetings and the end of most travel have made meeting platforms, such as Zoom and Microsoft Teams, 'must haves' for organisations and individuals, with the expectation of universal, end-to-end encryption

on servers in explicitly disclosed and widely trusted geographies. In rich countries, telehealth has increased in popularity, as have Internet-enabled devices to monitor personal health and fitness, with anxiety over how secure the video conversation with a health provider is, and where the video or personal data might end up.

Surveillance and tracking software from an Israeli defence contractor exporter, intended under its export licence only to combat terrorism and child trafficking, had been planted on telephones of journalists, dissidents, and opposition politicians in more than a handful of democracies. That no US-based numbers could be tracked, and how quickly the story disappeared, calls into question the whole expectation of 'impenetrability' that users should have of encryption. Apple's delay of the roll-out of ex ante, client-side scrutiny of users' phones for child pornography should at least temporarily assuage concerns that the company will no longer be able to resist search warrants issued for national security and law enforcement purposes; they *can* do it, so what can they be required, legally, to do? Breaking device encryption also violates user trust and device security assumptions, even if individual privacy and security concerns vary among users and across jurisdictions. Any system designed to detect data stored on a phone could also be used against activists, dissidents, and minorities by authoritarian regimes.

Social media platforms have been accused, with ample evidence, of enabling users to knowingly and deliberately spread disinformation that has endangered global public health, and thus complicated government efforts to bring vaccinations to their citizens. Elections in the US and France were attacked by state or state-sanctioned actors, spreading disinformation and rumours. The algorithms that cleverly feed targeted advertising have made Alphabet and Facebook management and shareholders very wealthy. They have not so far been effectively and consistently deployed to detect and remove disinformation, even if the twelve main sources of vaccination falsehoods are known[1]. Nor do the companies seem to combat misinformation so that it no longer spreads freely among platform users sorted by their 'likes and shares' into information-rich and misled communities, rather than just on their shared appeal for cat pictures, recipes, and internet shops.

Expectations that governments will bring this under control are misplaced. The EU's General Data Protection Regulation (GDPR) sets the current gold standard for the collection and protection of personal data, but nowhere else in the democratic world is similarly comprehensive and enforced regulation in place. Legislators in the US are arguing that repeal of Section 230 of the Communications Decency Act to make platforms responsible for what users post has been successfully resisted and may not pass constitutional muster. Other countries, even a few democracies, have pressured social media providers to remove material critical of political leaders or have legislation, for example, that prohibits online Holocaust denial.

And then there is China.

The China Factor

The Great Firewall and its army of moderators have limited political debate and even searches, while advances in AI have expanded its reach and effectiveness. More recently, though, China's pervasive deployment of facial recognition hardware and software, and advanced machine

learning using data from domestic databases and platforms and purloined foreign data sources like personal information stolen from the US's Office of Management and Budget, leave many concerned. The security threat from Huawei's 5G devices remains a point of contention among the US and its allies. Meanwhile, China's taking advantage of a software flaw to hack Microsoft Exchange cloud servers left organisations worldwide wondering about the reliability of the cloud concept itself even if email and other databases were not penetrated in this case[2].

China is not alone, though. Two disgruntled ex-journalists recently conducted open-source investigations using purchased cell phone location data and other information to identify a senior American Catholic priest and tie his location to use of Grindr, a gay dating app[3]. But crowdsourcing using openly available data from social media platforms also led to the identification and arrest of over 600 Donald Trump supporters who on 6 January 2021 attempted a violent coup against the US government and the peaceful transfer of power.

Will the "mutually assured surveillance" of open-source intelligence, freely accessible data, and the tools to analyse it demystify the world and make it less dangerous, or will the freedom of data erode personal privacy, and trust, on which free societies depend[4]? Will the detection of flaws in control systems, including AI-enabled expert support, promote freedom, or make everyone suspicious and clamouring for secrecy?

So where do these interrelated cyber issues fit together, and what is the way forward?

Is Change Coming?

It is no secret that the tech giants see regulation as imminent and affecting them and their lucrative business models; new laws and regulations will affect their users too. They will all lawyer up. Tech giants will wisely ally with the lawyers of their customers, and once the US Congress turns its polarised eyes to them, lobbyists will be agitated. The two major app stores will be a battleground, too, with competition advocates demanding open access, existing actors preaching the security advantages of their commercially walled gardens, and most app developers and app downloaders not feeling strongly except when something massive goes wrong. New South Korean legislation cracking open commissions paid by app developers is being admired in Europe and the US[5]. Gaming platforms seem to have figured out how to operate and be profitable without casting their dice (or arming their warriors) in non-compatible ways.

Device makers will need to address the Pegasus problem and strengthen device-level security against malware that can be placed on users' devices without any user action or awareness. Cloud service providers and internet service providers will need to reinforce and publicise steps they have taken to identify, attribute, and neutralise malware that their systems were used to transmit.

The US and the EU should work together to make this happen. India and China—active outposts of the major global software and platform companies, talented device designers and aspiring manufacturers, major, successful IT service providers, and with significant user populations—need to buy into this effort. India, as an IT player in many areas, could perhaps be an honest broker on malware that could affect critical infrastructure, providing quality assurance and seeing a national

interest in encouraging protection of critical infrastructure (of which India has and needs much more). India's hands are, for the moment, relatively clean on cyber-espionage, and as a victim of disinformation but not yet a locus for malevolent non-state behaviour or cybertheft of other countries' intellectual property, its active engagement is to be welcomed and is surely coming.

The pandemic has disrupted global supply chains, including for relatively lower value chips used for automobiles and consumer electronics. This would be the time to examine quality assurance of ubiquitous internet-of-things chips to ensure they do what their customers build their complicated devices to do, that their design is both easily reprogrammable and resistant to cyber tampering. Autonomous electric vehicles and robotic medical devices have captured the 5G narrative. They must work as designed, even if neither the devices nor the well-known, licenced, regulated, and broadly trusted human activities (driving and surgery in this case) they will enhance and enable are error-free. Neither robotic device was instructed how to drive, or did residency beside a skilled surgeon, to develop rapid, risk-based judgement.

AI and machine learning pose huge challenges. Algorithmic bias is well documented, with contentious arguments about data collection, how the machines learn (or learn from each other), and how the privacy and security undertakings to users apply across databases that are in or leak into the public domain. On the one hand, seeing invisible patterns can inform discussion about social and racial justice, including investigative journalism by reputable news organisations, now an important feature of the public square. Data patterns analysed and interpreted across sources also advance knowledge and practical applications in long-quantified fields like medicine and engineering, and newcomers with "physics envy" in the social sciences. On the other hand, crowd preferences for data sources and collections, for algorithm-based analytical tools, and for preferred channels of dissemination, may exacerbate bias and misinformation in online communities. They may also taint broader search engine and information channels that rely on well-known platform features, whether on social media or search engines. 'Popular' isn't always 'true'.

As important is deciding how autonomous AI will be allowed to act based on its own analysis, whether a robot doing remote surgery, a driverless vehicle swerving to avoid an obstacle it may not have seen before (or may not correctly identify, like a Tesla approaching a parked emergency vehicle), or an armed drone looking for vehicle or group behaviour outside externally-defined norms of 'harmless'.

Some of this is the stuff of science fiction, the best technology being indistinguishable from magic. Life and politics do not operate under simple, universally accepted laws, however, and nations see important commercial, trade and investment advantages in getting as much of what they want as they can.

Relatively simple fixes in technology and regulation of what we know needs improvement, and transparency in what is being done and who is accountable for results and verification, are important first steps. Work to create trust-based relationships to tackle the more complex and the unanticipated issues needs to begin. Governments need to get started.

Looking into the crystal ball about an evolving ecosystem, there are three areas with particularly global risk and challenge.

First, as technology advances, and innovators, manufacturers, vendors and users lust after the newest, smartest and most advanced 'stuff', who is responsible for unevenly coded, increasingly bulky operating systems, zero-day attacks resulting from badly designed software and apps, or chips and hardware with manufacturing flaws acquired and assembled in a complicated global supply chain? Should the vendor of 'smart home' devices (some critical, some not) be on the hook legally for a bad shipment of embedded microchips, particularly those that connect wirelessly to the internet? How does the manufacturer of a device, probably purchased online and maybe installed by the user at home, replace the flawed component and how does the manufacturer of the flawed, probably low value basic chip face accountability, "make repairs" and rebuild its credibility as a supplier?

Consumers must beware, and suppliers will be lawyering up, and not just relying on their designers and engineers. Clearly governments play a traditional and long-accepted role to protect their citizens and consumers, and to set and then oversee some minimum standards of safety. This is much as electrical devices and industrial products need to meet international standards once they cross borders or appear on stores' shelves. What is the UL, CE or CSA equivalent for microchips and software, what are the criteria for gaining and retaining such certification, and who sets a portable, agreed and enforced standard? Or are operating systems, apps, infrastructure (including servers for the cloud) and devices themselves such sources of competitive advantage that 'makers' will resist or even fight certification of beyond the marketplace and, when mishaps occur, regulators act? Will the actions of one country, on one vendor of hardware, software or services, be welcomed, challenged, resisted or defied in other jurisdictions?

Second, will the protection of data, in particular personally-identifiable data, within and across jurisdictions, align to some global standard? The US, EU and China have all acted recently, and face pressure to do more. Powerful technology companies have faced huge fines (in the EU under GDPR), regulatory punishment and threatened legislation or break-up (in the US's highly polarised polity), and limitations on external financing and their scope of non-tech activities (in China, for stated fear of data loss as the power tech platforms seek to list on American stock exchanges). These reflect the innovation of key tech actors in the US and China, the data protection leadership in the EU, and—probably—national security concerns in China about their private companies' reach into individual and family financial, social media and consumer behaviour, in competition with the state's own data collection, including biometric information, in a surveillance-based administrative state. Even if the US and the EU enforce new standards of data retention and data security, data protection will need to be reinforced against both domestic and foreign snooping and tampering, whether for profit, intellectual property theft, political meddling, or espionage in a new virtual environment. And if China's goals are to have a gated garden of its own data, does its nativist protectionism inhibit its own tech innovators from seeking international markets and opportunities, where some measure of competition might well benefit consumers by increasing the choices they face and bringing them new, exciting services and hardware they use to access and enjoy it?

The third, and highly speculative possibility—where the crystal ball remains very cloudy—is how this all comes together. Are Facebook and other companies right when they say virtual reality (VR) is our future, and not just in it? Will a *metaverse*, an immersive digital environment that is not VR glasses or headsets, become a 'thing', much as social platforms have already created

new jobs like influencers and new ways of social interaction? Will today's highly competitive and rapidly evolving tech marketplace give way to a collection of technologies, from the hardware to the wiring, that will work seamlessly together? Will this new ecosystem defy expectations about privacy, and challenge regulators to regulate the embedded AI and rid it of algorithmic bias, and catalogue, regulate and oversee the huge databases that will learn and interact to make it happen?

Unlike a physical universe operating according to the laws of science, what social and ethical principles would underpin a new, technology-enabled ecosystem to make it sustainable, inclusive, and fair, and keep it that way?

Endnotes

1. Elizabeth Dwoskin, "Misinformation on Facebook got six times more clicks than factual news during the 2020 election, study says," *Washington Post*, September 4, 2021, https://www.washingtonpost.com/technology/2021/09/03/facebook-misinformation-nyu-study/.

2. Eric Tucker, "Microsoft Exchange hack caused by China, US and allies say", *AP News*, July 20, 2021, https://apnews.com/article/microsoft-exchange-hack-biden-china-d533f5361cbc3374fdea58d3fb059f35.

3. Liam Stack, "Catholic Officials on Edge After Reports of Priests Using Grindr", *New York Times*, August 20, 2021,https://www.nytimes.com/2021/08/20/nyregion/pillar-grindr-catholic-church.html.

4. "The promise of open-source intelligence", *The Economist*, August 7, 2021, https://www.economist.com/leaders/2021/08/07/the-promise-of-open-source-intelligence.

5. Mitchell Clark and Jon Porter, "Apple and Google must allow developers to use other payment systems, new Korean law declares", *The Verge*, August 31, 2021, https://www.theverge.com/2021/8/31/22643800/apple-google-south-korea-app-store-payment-legislation-passes.

Bridging the Space Governance Divide: Beyond East vs West Dynamics

Daniel Porras

Few issues create as clear a divide among United Nations (UN) member states as that of the Prevention of an Arms Race in Outer Space (PAROS). Since the UN first took up this issue 40 years ago, states have consistently voted along lines that mirror geopolitical divisions present during the Cold War. As a result, despite a growing recognition throughout the world that the status quo is not sustainable for stability and security in space, neither the Conference on Disarmament nor the UN General Assembly have managed to reach a consensus on any form of space security agreements.

Over the years, there were numerous examples of space security discussions that broke down along these divisive lines. Yet one of the most contentious issues continues to be the debate over whether the UN should adopt an instrument that is politically or legally binding. On the one hand, a group of states that includes China, Russia and most of the Global South say that the issue of weaponisation of outer space is of such importance that it merits a new treaty. On the other hand, a group of states that includes the US and most of its major allies insist that the issue is too imminent for a treaty, and so political norms should be pursued.

Every major discussion on space arms control over the last decade has faced this fundamental question, from the EU's draft 'International Code of Conduct for Space Activities' (ICoC) to the most recent Group of Governmental Experts (GGE) on PAROS. While the subject matter of these initiatives has varied, the debate over treaties and political norms are always present. Likewise, it will be present this fall when the UN debates whether it should form an Open-Ended Working Group on reducing threats in space through norms of behaviour. As such, states will need to find a bridge across this divide no matter what the subject matter of any future discussions on space security. Otherwise, no initiative will be able to reach the critical mass of adherence necessary to make it worthwhile. The hope is that they can find a compromise before either side is proven right.

State of Space Security

Until 2007, only two countries had ever demonstrated a desire to obtain counterspace capabilities—the ability to deny, degrade or even destroy a space object. While the US and the Soviets had actively tested several designs throughout the Cold War—including nuclear weapons, direct ascent missiles and co-orbital satellite destroyers—no serious developments emerged following the fall of the Soviet Union. However, in 2007, China became the third country to successfully test a kinetic direct-ascent anti-satellite weapon (ASAT), at an altitude of 980 km. This test became a signal to the rest of the world that more than one country could now target and destroy a satellite. Moreover, three of the world's major nuclear rivals now possessed the means to undermine an opponent's ability to deliver a nuclear strike. While experts and diplomats had discussed such a possibility for several decades, the Chinese ASAT test marked a turning point in how real threats to satellites had become.

Since then, there are numerous indicators that the world is entering an arms race, and space capabilities are a part of it[1]. First, geopolitical rivalries are at their worst since the end of the Cold War, notably between China, Russia and the US. Second, rivals are developing corresponding capabilities to interfere with space systems, such as direct-ascent ASATs or co-orbital vehicles. Third, the development and deployment of specialised military units, as well as counterspace capabilities, signals an acceleration in the development of space-related armaments. As such, militaries are positioning themselves more and more to use the denial of space capabilities as options for future conflicts.

Today, the consequences of open conflict in space are uncertain, but they do pose a real threat to the long-term sustainability of outer space. As such, UN member states are actively seeking the means to mitigate or prevent a conflict that involves space objects. Unfortunately, this recognition of a problem is usually as far as the UN gets before it starts debating over whether to have a treaty, with legal obligations or to have political agreements, which are non-binding. One side argues that only a treaty is strong enough to ensure the long-term sustainability of outer space, while the other argues that the issue is so dire that only political norms can be adopted in time to avoid a catastrophe. The UN has run into this debate for decades, and there are no signs yet that any new discussions will not continue to do so.

Early Days of PAROS

The initial discussions around PAROS date back to the early 1980s during the 10th Special Session on Disarmament. The first two resolutions that emerged already signalled a difference of opinions in what PAROS meant. The first resolution (A/RES/36/97C), sponsored by the Western European and Other Group (WEOG), sought "an effective and verifiable agreement to prohibit anti-satellite systems". This proposal included elements such as "anti-satellite systems designed to impair the functioning of, interfere with, damage or destroy satellites of other nations"[2]. The second resolution (A/RES/36/99), put forward by the Eastern European and other states, focused on the prohibition of placing weapons in outer space, including conventional ones. These two positions can be simplified by noting that one of the resolutions focused on threats *to* space systems while the second focused on threats *from* space systems.

To bring these two positions together, member states created an ad hoc committee, but progress was slowed by a familiar debate over its mandate. The Group of Eastern States wanted the committee to be a formal setting for negotiations on a treaty, while the Eastern States only wanted it to be of an exploratory nature, with formal negotiations to come later[3]. In the end, the committee was formed to examine substantive and general considerations but only as a first step towards PAROS.

The committee met for ten years and, in 1994, a set of consultations sought to answer, among other things, whether or not the existing space treaties were sufficient to ensure PAROS, or whether there were indeed gaps in the law that required new legal instruments[4]. Here, there was a considerable difference of opinion. A group of 21, led by China, argued that the existing legal regime (which focused exclusively on nuclear weapons and weapons of mass destruction) was not sufficient, leaving open the possibility for the deployment of conventional weapons, high-energy lasers, and particle beams[5]. They also expressed concern over the development of

space-based missile defence systems that would double as anti-satellite systems. However, the WEOG states, notably the US, argued that the existing regime was sufficient to ensure that space would continue to be used for peaceful purposes, questioning whether there even was an arms race in space[6].

Modern Divisions

Over the next few years, China continued to argue that the existing legal regime was insufficient to ensure PAROS, and made several proposals for a new instrument[7]. Whilst it was unable to make any traction with the Western states, it did find considerable interest from much of the rest of the world. Then, in 2008, following on from their own ASAT test in 2007, China and Russia proposed a Treaty on the Prevention of the Placement of Weapons in Outer Space (PPWT). This proposal sought to prohibit the placement of any kinds of weapons in orbit, as well as to further emphasise the prohibition of the threat or use of force against objects in space. The reception of this proposal split down familiar lines.

In its response, the US reaffirmed its position that, first, the proposal was not needed as the existing regime was sufficient and, second, that the proposal was flawed[8]. Notably, the definition of weapon was not workable due to the dual-use (multi-use) nature of space activities, and because it was not effectively verifiable[9]. This position, shared by most US allies, has not changed to date.

The debate over political versus legal agreements became even more prominent during the open-ended consultations over the EU's proposed ICoC (2012-2015). This effort, held outside of the UN system, sought to establish voluntary norms of responsible behaviour for space activities. It contained several proposed politically binding norms on issues such as space debris, notifications, and even the application of international humanitarian law to space. Despite many countries agreeing on most of the substantive or technical issues, the EU was unable to bridge several diplomatic hurdles, including that of political vs legal agreements[10]. For most countries, including the BRICS nations, adopting a legally binding instrument was not only urgent but a priority over adopting mere political norms. However, the position of the Western allies remained that the only option was political norms. In the end, the two sides could not agree, and the effort lost momentum.

In 2017, space security once again emerged in UN discussions. This time, the Russian and Chinese governments proposed a GGE on PAROS, with the mandate of examining possible elements of a legally binding instrument (A/RES/72/250). This resolution was adopted by a vote of 108 in favour to five against (France, Israel, Ukraine, UK, US), with 47 abstentions (including nearly all of Europe). Despite concerns that some of the key spacefaring nations would not participate in this GGE, in the end, most of the major Western players were involved, including the US and many of its allies. Simultaneously, the Conference on Disarmament formed subsidiary bodies to examine the different items on its agenda, including PAROS. Ambassador G. Patriota of Brazil chaired both groups. In the end, Subsidiary Body 3 was able to adopt a report (CD/WP.611) while the GGE could not. However, in both discussions, the debate over politically versus legally binding agreements was present, and the debates did not come any closer to reaching a compromise.

New Opportunities

Following the GGE on PAROS, the UK government launched another initiative (A/RES/75/36), adopted in the UN General Assembly in December 2020. This resolution did not put forward a concrete solution but instead acted as a means to launch a dialogue. It asked the UN Secretary-General to carry out a survey that sought to determine what are space threats, what can be norms, and what could be the way forward. Thirty countries and the EU submitted responses, as well as several intergovernmental and non-governmental organisations[11]. Numerous responses contained similar views, with several countries not usually aligned showing convergence on issues like ASAT testing. However, one issue that remains evident is that of legally versus politically binding agreements. The UK has already announced that it will be submitting a follow-up proposal to form an Open-Ended Working Group to examine the development of norms for space security. It is almost certain that the debate over legal versus political norms will once again come up and, if not handled appropriately, could lead once more to a deadlock. Meanwhile, behaviours that put space security at risk will continue to proliferate.

A Bridge Not So Far

Over the last ten years, all measures passed in the UN around space security have been by wide margins. Whether they deal with political or legally binding agreements, most countries are in favour of taking some form of action to strengthen stability in space. Indeed, the failure of the EU-ICoC was not only due to the EU's unwillingness to discuss legally binding agreements, but this debate certainly exemplified what many perceived to be the West's "take it or leave it" approach to space security. Throughout the process, many states expressed a willingness to go further than what the EU sought. In the end, they were left to wonder: if this issue is so serious, why not adopt a treaty?

The PPWT has been criticised for not being workable, but that does not mean that the UN cannot pursue other options, which are effectively verifiable. States have proposed such an idea, including a ban on the live testing of kinetic ASAT weapons. However, the possibility of a treaty continues to be something of a taboo among Western countries.

One option that could be a compromise is to recognise the development of political norms as a first step towards creating the conditions to one day have a treaty. In this way, states could begin to develop trust in each other through small, focussed steps. Likewise, states could develop trust in the process by adopting widely supported measures that all space actors are likely to follow. Seeing that simple measures are possible, states could work on increasingly more ambitious measures. The goal of each of these measures could be to create a political and technical environment where the adoption of a treaty over some to-be-determined subject matter is possible. By so doing, Western countries can signal to the rest of the world that, at the very least, they respect the desire to have a treaty and that they are willing to compromise. Such a concession could be key in getting past one of the most basic political hurdles that has hampered PAROS discussions for decades.

This idea of adopting political norms as a first step could be formally embodied in a norm that emerges from discussions at the UN. In the preamble of a relevant resolution, language could be inserted that says something to the effect of: 'with a view towards creating the conditions to one day be able to adopt a treaty on PAROS…'

This language will not bind any parties to any timeline or subject matter discussion but will be sufficiently concrete to give international partners a positive signal that their views are being taken seriously.

Conclusion

The debate over politically versus legally binding agreements has sunk a number of discussions on PAROS, and it will likely continue to do so unless it is handled diplomatically. One option to sidestep this hurdle is to formally acknowledge that many countries see this challenge as meriting a new legal agreement, one that will hold all parties accountable. One way to show this acknowledgement is to recognise the development of politically binding norms as a first step towards creating the conditions for a treaty. By so doing, Western countries could secure the support of the Global South, winning so much political backing that reluctant space powers will

Endnotes

1. Benjamin Silverstein, Daniel Porras, John Borrie, "Alternative Approaches and Indicators of an Arms Race in Outer Space," United Nations Institute for Disarmament Research, May 2020, https://unidir.org/publication/alternative-approaches-and-indicators-prevention-arms-race-outer-space

2. United Nations, "Yearbook of the United Nations 1981," UN, p. 81, https://www.unmultimedia.org/searchers/yearbook/page_un2.jsp?volume=1981&page=1

3. Paul Meyer, "The CD and PAROS: A Short History," United Nations Institute for Disarmament Research, April 2011, p.2, https://www.unidir.org/files/publications/pdfs/the-conference-on-disarmament-and-the-prevention-of-an-arms-race-in-outer-space-370.pdf

4. It is worth noting that, while many delegations found the lack of established terminology to be an issue, the Italian Friend of the Chair concluded that having legal terms was not a pre-condition for the adoption of new measures on PAROS; United Nations Digital Library, "Conference on Disarmament, Report of the Ad Hoc Committee on Prevention of an Arms Race in Outer Space," August 24, 1994, https://digitallibrary.un.org/record/188861?ln=en

5. "Conference on Disarmament, Report of the Ad Hoc Committee on Prevention of an Arms Race in Outer Space"

6. "Conference on Disarmament, Report of the Ad Hoc Committee on Prevention of an Arms Race in Outer Space"

7. Conference on Disarmament, "Working Paper: China's Position on and Suggestions for Ways to Address the Issue of Prevention of an Arms Race in Outer Space at the Conference on Disarmament," CD/1606, February 9, 2000, https://undocs.org/pdf?symbol=en/

likely have to adopt any emerging norms from this process.

CD/1606; Conference on Disarmament, "Working Paper Presented by the Delegations Of China, The Russian Federation, Vietnam, Indonesia, Belarus, Zimbabwe And Syrian Arab Republic: Possible Elements for a Future International Legal Agreement on the Prevention of the Deployment of Weapons in Outer Space, the Threat or Use of Force Against Outer Space Objects," CD/1679, June 28, 2002, https://undocs.org/pdf?symbol=en/CD/1679

8. Conference on Disarmament, "Analysis of a Draft "Treaty on Prevention of the Placement of Weapons in Outer Space, or the Threat or Use of Force Against Outer Space Objects"," CD/1847, August 26, 2008, https://undocs.org/pdf?symbol=en/CD/1679

9. Daniel Porras, "Eyes on the Sky," United Nations Institute for Disarmament Research, https://unidir.org/publication/eyes-sky

10. Lucia Marte, "Code of conduct on space activities: unsolved critiques and the question of its identity,", Fondation pour la Recherché Strategique, Note de la FRS n°26/2015, December 17, 2015, https://www.frstrategie.org/en/publications/notes/code-conduct-space-activities-unsolved-critiques-and-question-its-identity-2015.

11. United Nations Office of Disarmament Affairs, "Report of the Secretary-General on reducing space threats through norms, rules and principles of responsible behaviors," 2021, https://www.un.org/disarmament/topics/outerspace-sg-report-outer-space-2021/.

Solutions

India's Opportunity to Lead by Example in Collaborative Technological Evolution

Nisha Holla

Global technology giants, overwhelmingly American or (lately) Chinese in origin, were not created in a vacuum. The State, in both instances, has facilitated their development into multinational near-monopolistic corporations and is deeply embedded in them. The new policy challenge the US faces is that society has overwhelmingly declared that these companies loom too large, have become monopolies, and exert omni-directional influence on civilians. The country must figure out how to rein in the giants they nurtured. China, too, despite the already rigid state-controlled development agenda, is deploying tighter regulatory mechanisms to ensure its technology giants adhere to the boundaries being redefined. Governments worldwide are passing legislation to regulate the majoritarian control of foreign tech giants over digital territory and hand back agency to the ordinary user. India is also asserting its techno-sovereignty while creating indigenous technology giants. This is India's opportunity to make a resilient and inclusive environment, and repeat its success with digital public goods and India Stack, to set a new global standard in collaborative technological evolution and open-internet society governance.

State-Facilitated Technology Giant Creation in US and China

The US and China's mega-cap conglomerates have built multinational monopolistic empires and continue to grow in valuation, market share, and number of users, while enveloping emerging business models by acquiring adjacent platforms. The states in both the US and China have facilitated, directed, and become deeply embedded in their respective technological ecosystems. In both countries, public investment in innovation and research and development (R&D) is traditionally high priority, and only accelerating every decade.

Government bodies in the US deploy specific and time-tested strategies to facilitate technology development and create technological giants in every sector. One, it has built comprehensive public investment and procurement vehicles that encourage indigenous technology development and commercialisation. When public and private sectors are combined, the US utilises 3 percent of its US$22 trillion economy (pre-COVID-19) on R&D[1], or approximately US$660 billion annually. Every major government department has an extensive budget. The US Department of Defense, for example, has an annual budget of US$190 billion for research and procurement, and is the largest client of American defence technological companies like Lockheed Martin, Raytheon, Boeing, and Northrup Grumman[2]. These companies have been elevated to technological giants in their space, driven by multiplatform government contracts. Government R&D and procurement budgets generally enjoy bipartisan support, a rare consensus that indigenous technological development is crucial to maintaining socioeconomic leadership. The US's bipartisan imperative to actively procure from and engage the services of American companies has been essential in establishing its America-first technology doctrine. This model has been extended aggressively over digital procurement as well, from cloud capacity to cybersecurity.

Two, the US pioneers dual-use technology development and deployment models. Dual-use technologies are commercialised by the private sector and deployed for national use either by the government or the military. R&D funding is specifically allocated to develop dual-use technologies that benefit both the state and economic growth via the private sector. It also incentivises its technology companies to customise their platforms for the government's use cases, tightened with government-specific security, compliance, scope and controls, regulatory alignment, authentication, and encryption protocols. Amazon deploys a unique instance of its Amazon Web Services for the US government called AWS GovCloud. Alphabet's entity Google provides special services for the state, including Google Cloud, which assists the federal government in improving operational effectiveness and delivery of services with AI-driven analytics and data platforms, and Google Maps to improve public services and deliver critical utilities. Azure Government by Microsoft is a dedicated instance for US governments to host mission-critical applications. Multiple government bodies and military organisations have signed contracts with Palantir for customised data-driven platforms[3]. These are only some of the publicly known use cases of indigenously-developed US digital platforms. The State is an active steward of its technology companies with stable long-term contracts.

Three, the US has traditionally applied a light-touch policy vis-a vis the commercial models of its technology companies to actively facilitate their rapid growth and to reduce friction in the process. Launching companies, attracting investment capital, incentivising domestic and global investments, listing companies on the stock markets, and certifying products are straightforward processes in the US. India is only orienting towards these frameworks now. Accordingly, technology companies in that country easily attract seed and growth capital, proliferate, are supported by government contracts, and go on to list successfully on the US stock markets. Global technology giants like Facebook, Google (now Alphabet), Amazon, Microsoft, Twitter, Apple, Nvidia, Intel, Qualcomm and Cisco are products of this ecosystem that is engineered to create minimal policy friction to growth. Indeed, the abbreviation FAANG—Facebook, Apple, Amazon, Netflix, and Google—represents the fastest growing stocks in the US. Not only does it underscore the rapid growth that technology platforms make possible, but these firms collectively make up a whopping 19 percent of the S&P500, with a combined market capitalisation of over US$7 trillion (as of August 2021)[4].

The US State has actively facilitated the growth of its technology companies into multinational giants. Most countries' governments and their citizens utilise some form of American technology in their daily lives, especially digital platforms like Facebook (and its entities like WhatsApp and Instagram), Twitter, Netflix, Amazon, and a host of Google applications like Mail, Maps, Search, Drive, Photos, Play Store, Android, and Translate. American technology, primarily digital, has become ubiquitous today. Indeed, this is one of the ways the US has maintained its socioeconomic leadership in the world.

China took several leaves out of the US playbook and deployed a deeply State-driven strategy to build indigenous technology companies and propel them into multinational giants relatively faster. The State has built an equally massive public investment framework for technology development. China's innovation spending in 2020 was US$378 billion—2.4 percent of its GDP[5]. It actively incentivises its brightest to start technology companies with substantial equity grants, non-term R&D funding, and free or subsidised utilities. Within its communist economic

framework, the Chinese State has developed several capitalist strategies to attract large pools of global capital to invest in its companies via its special economic zones, hi-tech development zones and open coastal cities. Banning American digital companies like Facebook and Google, combined with a billion-strong market hungry for new digital products and services has created a pull-effect for indigenous technology companies to fill the vacuum. These companies started providing services their American counterparts could no more but quickly extended into adjacent services, actively encouraged by the State. This has created State-blessed tech oligarchies that are expanding globally, listing on foreign capital markets, and competing heavily with American technology giants to gain a global market share.

As of 2021, China has over 170 unicorns[6], and technology companies dominate economic growth in the country. Companies like Tencent, Alibaba, JD, Huawei, Baidu, BBK Electronics, and Xiaomi are large multinational conglomerates that rank among the world's most highly profitable companies. Their reach and digital networks rival US companies, and they have succeeded in building significant global followings with their digital and electronic offerings, often cheaper than American and other Asian equivalents.

Big Tech vs. the Government

Big Tech is now clashing with governments around the world, democratic or otherwise. While the details of how governments grapple with Big Tech's overwhelming envelope on citizens and modern society varies across regions and systems, it is reasonably evident that the public pursuit of balance and regulatory adherence is a clumsy and stepwise walk in a maze of bureaucratic and technocratic jargon.

The European Union (EU) was the first region to take a hard-line stance on data privacy, announcing the General Data Protection Regulation (GDPR) in 2018. After two years of public posturing and verbal flagellation, the EU had only issued an underwhelming number of fines to two US tech companies as of May 2020[7]. A loophole in the GDPR meant that only the Irish Data Protection Commission could handle legal cases involving cross-border data complaints since most tech companies were domiciled in Ireland for tax purposes. Under the stated "one-stop-shop" mechanism, even if another EU member state proposed to sue a US tech company for violating privacy rights, they would need the Irish authority to adjudicate on the case[8]. While a recent EU Court of Justice ruling has concluded that tech companies can now be sued under certain conditions by any data protection authority in the EU, this may just make the enforcement of GDPR that much more fragmented and inconsistent.

The US government has pursued a different approach to reining in their mega-cap technology conglomerates. In October 2020, a highly partisan Senate hearing cross-examined the founders and CEOs of the largest US tech companies in vastly differing ways. While the Democrats were insistent on tech companies taking more accountability for the control of content and communication on their platforms, the Republicans expressed deep unease at the high degrees of politically or ideologically-inspired censorship enforced on the same platforms[9]. These contradictory signals made for entertaining watching for the rest of the world but certainly exposed the chasm in the alignment of principles of free speech, censorship, the State's role, the nature of outdated regulation and definitions, and the difficulty in reaching a consensus around a common-sense framework.

In June 2021, US lawmakers of the Biden administration introduced multiple bills that attempted to address topics like data, mergers, and the anti-competitive behaviour of tech companies[10]. "Big Tech's unchecked growth and dominance have led to incredible abuses of power that have hurt consumers, workers, small businesses and innovation," said Robert Weisman, president of the advocacy group Public Citizen. However, criticism that these laws target specific companies and not business practices is valid and will lead to further delays in implementing these frameworks.

Across the Pacific, China's action against its extensive tech ecosystem has been sharp, targeted, and swift. Over US$1 trillion of market capitalisation has been destroyed after the shock ban by the Chinese government of different business models and planned IPOs[11]. For example, China has mandated that all education technology companies be required to convert into non-profit entities, thereby killing the business models of a thriving tech segment in the country. Further, Didi Global lost over US$22 billion in market capitalisation in July 2021[12] after the Chinese regulator expressed concerns about the company's handling of citizen's data and enforced an immediate delisting of the company's apps from the country's app stores. This event came as a surprise to global investors, who had just bought into Didi's upsized US IPO in June 2021. Over just a few quarters, the Chinese State has systematically shifted policy after decades of committed support and is altering the fundamentals of several technology sectors.

Technological giants in the US were not created in isolation either. The State has played a critical role in creating so many mega-cap companies, prolifically and consistently. The unyielding commitment to the America-first doctrine is bipartisan, top-down, and consistent with an extremely long view of maintaining leadership. It maintains trust in its citizens' capabilities to lead the global technological edge over decades in multiple frontiers. The State plays an indispensable role by allocating long-term investment, being an early adopter and placing large-scale procurement contracts but otherwise does not micromanage. With this methodology, it is unsurprising that the US State has wielded its dual-use mechanism to create tech giants in every critical industry—internet, communications, semiconductors, defence, aviation, robotics, pharmaceutical, cybersecurity, material sciences, and space. The State is amongst the early adopters and remains the biggest client for a multitude of firms, from Google and Microsoft to Amazon AWS and Palantir; the State has become deeply embedded into these tech monopolies. This system has served American interests, both domestically and globally, exceedingly well.

The new policy challenge facing the US State is that society has openly declared that these companies loom too large, have grown into monopolies and exert omnipresent influence on civilians. The country must figure out how to rein in the giants they nurtured while keeping their companies at the pinnacle—powerful and propelling an America-first tech doctrine while not threatening their own population. It is a bipartisan problem, a challenge that the US government must solve regardless of who is in power. In this regard, the US and China have a similar problem.

China's determination to dominate the technological edge is State-facilitated and directed. Since liberalising its economy in 1978 and its official World Trade Organization inclusion in 2001, China has developed innovative and attractive models to invite global companies, particularly American, to develop intellectual property and manufacturing bases in China. It has actively incentivised global capital to invest in its tech companies to create State-blessed tech oligarchies and walled-garden competition. With more active State facilitation than the US, China achieved

in 25 years what the US did in 50. With over 170 unicorns, China has taken a different stance on redefining what it considers unacceptable beyond its installed boundary conditions.

India's Relationship with Big Tech

India has taken neither the American nor the Chinese path actively in its commitment to technological resiliency. The overarching role of tech giants in society is an issue when your country has created them. India is still in the process of giant creation, and must take lessons from both the US and China and do it in 10 years and differently. Creating giants are essential but doing it differently can mean India sets a new example for collaborative technological evolution for the benefit of its citizens.

India is among the most prominent internet markets globally, indubitably the largest open internet society, and crucial to the growth stories of global technology giants, especially American companies that have been closed off to the only other billion-strong market. Facebook has 410 million Indian users; its subsidiary WhatsApp surpasses it at 530 million, while Instagram has 210 million[13]. YouTube has nearly 450 million users in India, while Twitter has 175 million. Before being banned in June 2020, TikTok had 200 million users in India, its largest overseas market[14].

While the US and China try different methods to rein in their technological giants, India is trying to regulate global technological giants without having created any in the country. However, this is imminent since the startup ecosystem has favourable tailwinds in shaping them within this decade. India's startup ecosystem stands third in the world, after the US and China. The nation presently houses 55,000 startups with a combined valuation of US$315 billion[15]. Fifty-nine unicorns dominate the ecosystem with a combined valuation of US$180 billion, of which 20 were created in the last year alone[16]. Estimates suggest that by 2025, India could house 100,000 startups with a combined valuation of US$1 trillion and more than 150 unicorns with the current trendlines[17]. Technology is fuelling this rapid growth. The government has largely been a bystander in this process apart from driving India's digital public goods (DPGs), colloquially called the India Stack.

Meanwhile, the Indian government has already clashed with several global technology giants. Industry bodies and regulators have routinely flagged e-commerce companies like Amazon and Flipkart (now owned by Walmart) for unfair business practices and anti-competition behaviour. Their demand retention strategies, such as discounted flash sales and the promotion of their owned private brands in search results, have also evoked much resistance from India's small and medium enterprise lobby groups, who are usually merchants selling on these platforms. India's new draft e-commerce rules attempt to balance these important points of view while enforcing a higher standard for competition and supporting inclusive innovation[18].

The Indian government has also attempted to install common-sense frameworks for digital media ethics and data localisation and protection, amongst the first large democracies to do so. These rules expect technology companies, irrespective of country of origin, to house the data of Indian citizens in India, to enforce data privacy, to run accountable processes to address the grievances of users and creators, and to adhere to a consistent definition of ethics in the

interest of protecting free speech. Global companies like Twitter[19] and Mastercard[20] have already run afoul of some of these regulatory requirements. While the Indian government has sincerely attempted a light-touch and common-sense approach[21], it is interesting to note the undercurrent of a no-compromise approach towards protecting citizens' rights across these guidelines covering multiple sectors.

This is India's opportunity to extend its common-sense frameworks and build a resilient governance structure overseeing its relationship with global and upcoming domestic technology giants.

New Example for Collaborative Technological Evolution

India has already taken a radical approach to digital technologies by developing its DPGs in a public-private partnership model. India's DPG framework (India Stack) is designed as a multiplatform modular system with a core centred around three pillars—a comprehensive identifier system (Aadhaar), a revolution in the telecommunication industry resulting in widespread mobile proliferation, and the Jan Dhan financial inclusion programme. India Stack has powered the country's digital trajectory, driven financial inclusion and integration, and enabled unique governance and relief delivery models. Its open application programming interfaces, or APIs, have enabled the private sector to develop new models, platforms and services, accelerating India's digital and financial inclusion trajectory.

Open internet and digital parity must be founded on five techno-sovereignty-driven and democratic first principles—universal access; bias towards inclusion; inalienable rights (foremost being the rights to privacy, self-determination, security and personal safety, and not to profiled); recourse to the law; and frameworks that support continuous innovation on top of the technological networks[22]. India's DPG framework already adheres to or can accommodate all of these five democratic-first principles. Since these open-network DPG frameworks fuel India's startup ecosystem, the very foundation of technology giant creation in India has, intentionally or otherwise, been diametrically different from the US and China. It is democratic and controlled by common-sense rules.

The new Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules posted in February 2021 is evidence of this new arc. These guidelines, drafted collaboratively by the Ministry of Electronics and Information Technology, endeavour to empower ordinary users amid growing concerns around the lack of transparency and accountability, and diminished user agency. They cover social media networks, OTT platforms and digital media companies, and provide users with a clear grievance redressal mechanism[23]. The Personal Data Protection Bill proposed in 2018 is another example of India's thought leadership in handing agency back to its digital citizenry and upholding data sovereignty with data localisation measures [24]. When nations worldwide face the challenges of balancing many vital principles in their respective democracies, India's proposal reveals a balanced approach without imposing unreasonable restrictions on innovation and platform development essential to economic growth.

Proactive frameworks beat reactionary policies every time. Instead of periodically getting caught in a theatre of conflict with technology companies and multinational giants like governments worldwide, the Indian government can formulate a level regulatory playing field for domestic and global players alike. India has already demonstrated an appetite for developing platforms in the public interest and setting international standards for DPGs. It must extend this thought leadership into creating a resilient and inclusive environment for developing indigenous technology giants and repeat its success with India Stack to set a new global standard.

It is unyieldingly essential for India's techno-sovereignty to produce indigenous technological giants in every sector, including the digital space. Depending on global companies for critical digital services is akin to digital colonisation and antithetical to sovereignty. Indigenous giant creation should not become a casualty of this exercise in setting new regulatory mechanisms. The Indian approach is already distinctive and inclusive. In collaboration with private sector experts, the administration must examine how this regulatory framework will be different from global, somewhat unsuccessful attempts and what the Indian State can do more to protect its citizens while enjoying the benefits of nurturing its technological giants.

In the new normal post-COVID-19, citizens' digital security and integrity, data localisation, creator freedom of expression, and the sovereign's ability to enforce the rule of law to protect the nation's digital economy are inviolable requirements. The Indian government must approach these critical issues with an agile mindset to support and not add friction to the inevitable role digital innovation will play in the country's path towards becoming a US$5 trillion economy. India has a generational opportunity to set a new global standard for open internet society governance while encouraging its technology ecosystem. It is undoubtedly time for India to lead by example in collaborative technological evolution.

Endnotes

1. UNESCO UIS, United Nations, *http://uis.unesco.org/apps/visualisations/research-and-development-spending/*

2. Nature, "DARPA 'Lookalikes' Must Ground Their Dreams in Reality," *Nature News*, March 11, 2020, https://www.nature.com/articles/d41586-020-00690-5.

3. Vandita Jadejam "Palantir Technologies' Government Contracts Will Put It on Top," *Nasdaq*, June 28, 2021, https://www.nasdaq.com/articles/palantir-technologies-government-contracts-will-put-it-on-top-2021-06-28.

4. Jason Fernando, "What Are FAANG STOCKS?," *Investopedia*, August 19, 2021, https://www.investopedia.com/terms/f/faang-stocks.asp.

5. Sam Shead. "China's Spending on Research and Development Hits a Record $378 Billion," *CNBC*, March 1, 2021, https://www.cnbc.com/2021/03/01/chinas-spending-on-rd-hits-a-record-378-billion.html.

6. "Unicorns in China." *Tracxn*, September 1, 2021, https://tracxn.com/d/unicorn-corner/unicorns-list-china.

7. Katie Collins, "As the GDPR Turns 2, Big Tech Should Watch out for Big Sanctions," *CNET*, May 24, 2020, https://www.cnet.com/news/as-the-gdpr-turns-2-big-tech-should-watch-out-for-big-sanctions/.

8. "Big Tech Companies Exposed to EU Privacy Cases after Court Decision," *Euronews*, June 15, 2021, https://www.euronews.com/2021/06/15/big-tech-companies-exposed-to-privacy-challenges-after-eu-court-decision.

9. Tony Romm and Rachel Lerman, "Facebook, Google, Twitter CEOs Clash with Congress in Pre-Election Showdown," *The Washington Post*, October 28, 2020, https://www.washingtonpost.com/technology/2020/10/28/twitter-facebook-google-senate-hearing-live-updates/

10. Cody Godwin, "US Lawmakers Introduce Bills Targeting Big Tech," *BBC News*, June 12, 2021, https://www.bbc.com/news/technology-57450345.

11. "Investors Lose $1 Trillion in China's Wild Week of Market Shocks," *Bloomberg*, July 30, 2021, https://www.bloomberg.com/news/articles/2021-07-30/investors-lose-1-trillion-in-china-s-wild-week-of-market-shocks.

12. "Didi Loses $22 Billion in Market Cap after China Crackdown," *The Economic Times*, July 6, 2021, https://economictimes.indiatimes.com/markets/stocks/news/didi-loses-22-billion-in-market-cap-after-china-crackdown/articleshow/84171914.cms?from=mdr.

13. Press Information Bureau, Government of India, https://pib.gov.in/PressReleseDetailm.aspx?PRID=1700749.

14. Manish Singh, "TikTok Goes down in India, Its Biggest Overseas Market," *TechCrunch*, June 30, 2020, https://techcrunch.com/2020/06/30/tiktok-goes-down-in-india-its-biggest-overseas-market/.

15. Priyanka Iyer, "India to Have over 150 Unicorns by 2025, Startups to Employ 3.25 million People: Report," *MoneyControl*, August 9, 2021, https://www.moneycontrol.com/news/business/india-to-have-over-150-unicorns-by-2025-startups-to-employ-3-25-million-people-report-7295851.html

16. "India to Have over 150 Unicorns by 2025, Startups to Employ 3.25 million People: Report."

17. TV Mohandas Pai and 3one4 Capital, "India - A Startups Nation," *3one4 Capital*, August 30, 2021, https://www.slideshare.net/W-3one4/india-startups-nationtv-mohandas-pai-3one4-capital-aug-2021.

18. "New India E-Commerce Rules and Their Impact, Explained," *The Economic Times*, June 28, 2021, https://economictimes.indiatimes.com/tech/trendspotting/new-india-e-commerce-rules-and-their-impact-explained/articleshow/83914653.cms.

19. "Twitter Faceoff: Speculations over 'Twitter Ban' over Non-Compliance with Government Norms," *The Times of India*, May 25, 2021, https://timesofindia.indiatimes.com/india/twitter-faceoff-speculations-over-twitter-ban-over-non-compliance-with-government-norms/articleshow/82940266.cms.

20. Ashwini Manikandan, and Saloni Shukla, "Mastercard's Dual Record Maintenance Led to RBI Ban," *The Economic Time*s, July 22, 2021, https://economictimes.indiatimes.com/news/economy/finance/mastercards-dual-record-maintenance-led-to-rbi-ban/articleshow/84623039.cms?from=mdr.

21. TV Mohandas Pai, "The Digital Media Code Balances Many Priorities, Interests of Several Stakeholders," *The Indian Express*, February 26, 2021. https://indianexpress.com/article/opinion/columns/digital-media-ethics-code-platform-google-facebook-7205122/.

22. Nisha Holla, "Democratising Technology for the next Six Billion," *Observer Research Foundation*, October 19, 2020, https://www.orfonline.org/expert-speak/democratising-technology-next-six-billion/.

23. Press Information Bureau, Government of India, https://pib.gov.in/PressReleseDetailm.aspx?PRID=1700749.

24. Ministry of Electronic and Information Technology, Government of India, https://www.meity.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf

Open Data a Critical Tool for Crises: Can India Make Better Use of it?

Samuel Neufeld and Sridhar Ganapathy

After the initial outbreak of the COVID-19 pandemic, several governmental institutions across India created mobile applications and websites aimed at providing information on the availability of critical resources (such as hospital beds and ventilators) to the public. While caseloads were manageable, these platforms managed to provide information without significant error. However, when the pandemic began to resurge in April 2021, an increasing number of citizens began reporting that government data had become outdated or otherwise inaccurate[1]. Without access to real-time, verified, and aggregated information, hundreds of thousands of infected Indians and their families found the logistics of seeking medical care and resources heavily obstructed.

Unable to depend on official government sources, scores of grassroots networks sprang into action, broadcasting information as well as triaging and routing critical patients to available resources via social media and group-chats comprising friends, colleagues, and communities[2]. While testament to the tenacity and resourcefulness of Indian society, these ad-hoc efforts, executed in the midst of a crisis, were ultimately insufficient in overcoming what amounted to a population-level void of healthcare information.

This void is symptomatic of the larger problem plaguing the nation's data ecosystem—the availability of high-value public interest data (identified via the ten principles delineated by the National Data Sharing and Accountability Policy [3] and relevant to social welfare) is sporadic and its dissemination is not standardised, reliable or timely.

Open Data as a Tool for Emergency Response

Across government, civil society and the private sector, technology has increasingly been adopted to scale and optimise the delivery of public services, resulting in the collection of tremendous amounts of data. This data covers a wide variety of information—including the provision of social protection benefits, transit and mobility, climate, ecology, geospatial imaging and maps—that can provide a range of useful social, economic, epidemiological, and climatic insights. Since the early 2010s, Indian government agencies and several civil society organisations have made efforts to release this data to the public by creating data repositories, termed open data platforms, that are accessible via the internet. Indian open data platforms—such as the central government's flagship Open Government Data Platform[4], the Pune DataStore[5], the India Observatory[6], and the India Data Portal[7]—are useful in myriad sectors and contexts, including emergency response. The function-specific public health information platforms created as the pandemic unfolded underline this potential.

Unfortunately, the COVID-19 pandemic is not the last crisis situation that India will face. It is a matter of when, not if, the next hazard strikes, whether it be a climate-induced natural disaster or the outbreak of a new infectious disease. While these events are largely unavoidable, the resulting disasters are a product of human action and decision[8]. Open data platforms fuelled

by robust data management infrastructure, systems, and procedures—and established well before an emergency situation arises—can empower government, civil society, and the public to effectively confront a crisis by taking both anticipatory and real-time action. India can draw valuable insights from these examples.

Example 1: Enabling targeted, local action in urban environments

The initial spread and impact of the pandemic was largely concentrated in cities. In many geographies, these circumstances revealed cracks in the social protection measures available to urban populations. However, they also demonstrated the potential of open data to be harnessed in such environments.

In New York City, NYC Open Data [9]provides thousands of datasets produced by city agencies. During the pandemic, this information enabled residents to find local places to volunteer, choose more hygienic food delivery, and identify routes for walking that are amenable to social distancing, among other functions[10].

Example 2: Anticipating and responding to climate change and natural disasters

The threat of climate change, coupled with the human and economic costs of extreme weather events, make real-time disaster maps essential information for public authorities, civil society, and the public. This information can be used to inform decisions regarding public safety, such as regional evacuations, potentially affecting millions of people.

In Indonesia, one of the most natural disaster-vulnerable countries in the world, developers created PetaBencana.id[11], an open tool that scrapes information from users of social media platforms to produce real-time flooding maps. These maps have been used by residents, civic groups, and government agencies to assess the development of weather events and inform decision-making during emergencies.

Example 3: Monitoring the outbreaks of infectious disease

As the COVID-19 pandemic has shown, outbreaks of infectious disease can quickly spiral out of control. However, by creating sophisticated outbreak monitoring systems and opening up such data to the public, it is possible to halt the onset of an epidemic.

After the 2005 outbreak of dengue fever in Singapore, the government created what has become known as the 'Dengue Website'[12]—a publicly available, online mapping tool that enables users to identify the location of case clusters. By informing the population[13], the platform empowers individuals and institutions to better prevent and protect against infection.

Considerations for the Next Era of Open Data

Evidence that the pandemic has disproportionately affected vulnerable communities, both in India and abroad, has mounted with time[14]. Marked most starkly in our collective imagination by the grim photos of thousands of daily wagers returning from India's cities to their villages by

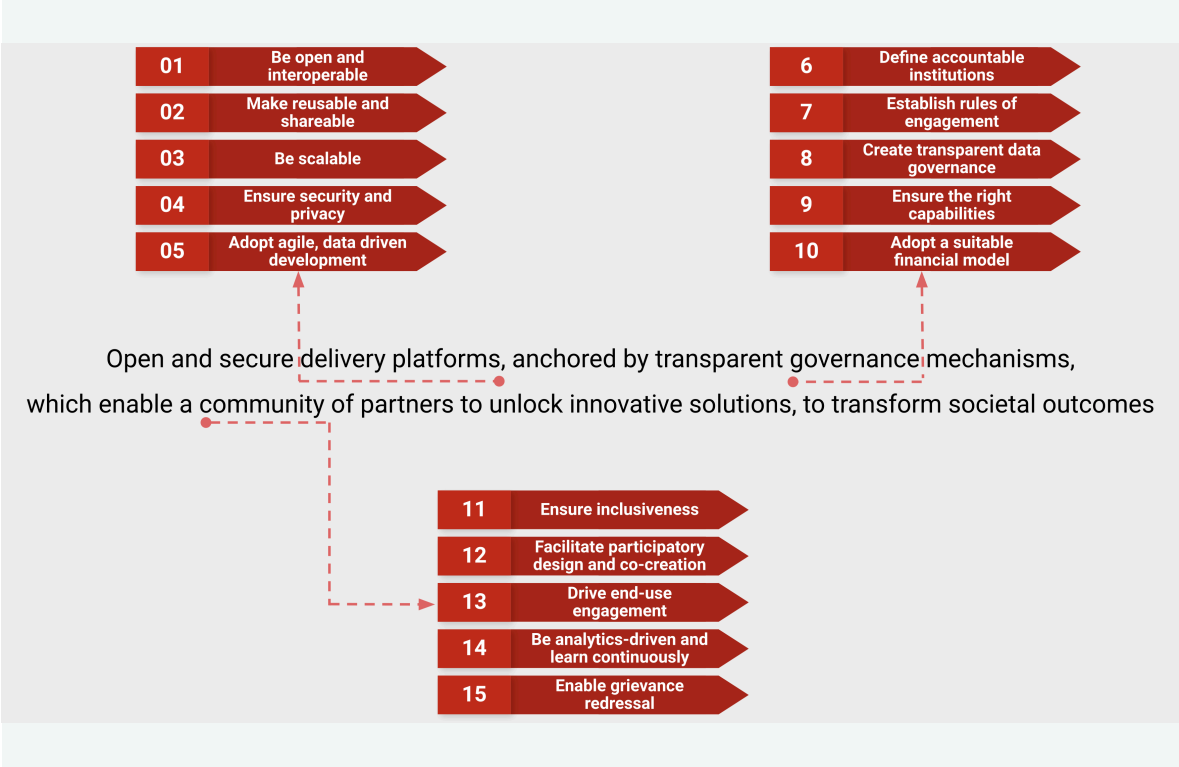
foot, marginalised communities have borne the brunt of COVID-19 and its socioeconomic and epidemiological impacts. In addition to those affected by the virus, millions have lost their jobs[15], slipped into poverty[16], or have had significantly restricted access to healthcare and education.

Meanwhile, the pandemic has accelerated the development and adoption of and reliance on digital technologies. Great strides in connectivity and cloud computing in the past decade have provided the technical bedrock to build and deliver apps catering to almost any requirement in near instant fashion, opening up a wide range of possibilities for use in governance and public service delivery. The inevitability of technological adoption across all facets of society—from working remotely to the provision of healthcare and education—poses increased importance on the value of data and raises time-sensitive questions about institutionalising guiding, ethical principles for its use. However, stark inequalities in digital access and competence threaten to leave behind millions of vulnerable people[17], and underline the urgent need for commensurate action to ensure data reflects gendered, ethnic, and community-centric considerations[18].

The accelerated adoption of digital technology in governance is poised to be a key enabler for open data in India. National policies for data sharing[19] and data governance assessments[20] are among the many initiatives consistent with the broader vision of using technology for societal outcomes, as delineated in the government's National Open Digital Ecosystems (NODE) strategy[21]. The NODE strategy features inclusiveness and participatory design as guiding principles, with a mission to reach and engage citizens through all channels while embracing security and privacy by design. User consent frameworks, encryption, and data collection limitations are to be utilised for such purposes. The common theme in such digital governance efforts is the focus on opening up data, enabling co-creation, fostering citizen participation, and spurring innovation.

Although the creation of human-centric open data platforms and the release of high-quality data will be one of the many outcomes of this strategy, realising the potential of open data will also require widespread adoption of a simple, open licence that maximises reuse[22]. While resources accessible on open data platforms are governed by the Government Open Data License[23], a wide range of public data is available on Indian government websites in the form of reports, tables and documents, which lack uniform guidance on reuse of such data. Not all data can be made public without significant effort, and in many cases, restricted access programs can be a good starting point for collaborating with the research community. Established models of secure data sharing exist for this purpose[24] and have proven they can deliver public benefit while protecting data confidentiality[25].

Figure 1: Guiding Principles for Building National Open Digital Ecosystems



Source: National Open Digital Ecosystems[26]

Roadmap for Strengthening the Indian Open Data Ecosystem

While the Indian government has made extensive efforts to provide the foundational policies and environment for open data to thrive, substantial reform is required to effectively leverage open data for time-sensitive functions like disaster relief. Indian policymakers, bureaucrats, and civil society leaders should consider pursuing the following objectives.

Maintaining adequate funding and skilled personnel

In many instances, even though relevant data platforms are operational, the institutions producing data suffer from chronic underfunding and insufficient human resources[27]. To execute complex tasks such as data collection, processing, engineering, curation, and publication, highly skilled human resources and ample bandwidth is needed. However, many governmental institutions and civil society organisations lack a strategic focus on data management, or struggle to obtain such resources with limited budgets.

Specialised personnel or teams, depending on the size of the organisation, are required for organisations to effectively manage and release data. These 'data cadres' should be supported with targeted government budget allocations at the national, state, and municipal levels. The proposal to initiate data and strategy units[28] within ministries/departments to create interdisciplinary teams covering programme monitoring, statistics, technology and analytics is a promising start. It specifically calls out the need to recruit personnel for the analytics roles,

which is lacking in the current structure of many government agencies and needs to be acted upon urgently. Moreover, supplemental support can be harnessed from grants and technical assistance provided by multilateral institutions and civil society groups dedicated to expanding the potential of data for social impact, like data.org.

Ultimately, to provide substantive value in the event of a crisis, it is critical that data collection and management, as well as the upkeep of open data platforms, receive ample budget and attention well in advance.

Unifying platforms through public-private-civic cooperation

Since the outbreak of the pandemic, more than 60 apps[29] have been launched by public authorities in India to help the public deal with COVID-19. However, many of them have overlapping functions.

Table 1: A selection of apps released by government authorities since March 2020

Function	Name	Developer	Geography
Contact Tracing	AAROGYA SETU	Ministry of Electronics and IT	Nationwide
	MAHAKAVACH	Maharashtra Government	Maharashtra
Quarantine Monitoring	BSAFE TRACKING	Kerala Police	Kerala
	SMC COVID-19 TRACKER	Surat Municipal Corporation	Surat
Healthcare FAQs	GOK DIRECT	Qkopy Online Services, Kozhikode; Kerala Government	Kerala
	KAVACH	Chhattisgarh Government	Chhattisgarh
Essential Services (Groceries, Medical, etc)	COVA PUNJAB	Uengage Services; Punjab Government	Punjab
	JAN SAHAYAK - HELPME APP	Haryana Government; OFB Tech	Haryana
COVID-19 Data Crowdsourcing	COVID-19 FEEDBACK	Ministry of Electronics and IT	Nationwide
	GCC CORONA MONITORING	Greater Chennai Corporation	Chennai

Source: The Print[30].

This duplication of efforts is not only extremely inefficient but also precipitates siloed repositories of data. In the absence of interlinked systems that can provide comprehensive information, members of the public often struggle to identify which platform has the specific data they require. A notable exception is Co-WIN, a scalable, open platform for vaccination). It has a dashboard[31] that provides comprehensive data about vaccination trends in the country, down to the vaccination centre level in each district. Valuable data, such as information about incidents of adverse effects, uptake in different age groups across time, and doses administered in urban and rural areas, are all available for public use.

In the context of India's federalised political architecture, its diverse demographic and socioeconomic composition, as well as the maturation of digital ecosystems in various sectors of the economy, the proliferation of function- or geography-specific open data initiatives is the need

of the hour. While attempts should be made to centralise efforts where appropriate, government authorities, technology companies and civil society groups should strive to focus their efforts on the creation of modular platforms that adhere to established standards of interoperability[32] and openness[33] to advance access to data among the general population. This approach also requires ensuring the availability of past versions of published data, an audit trail of changes, as well as the permanence of web-based resources[34]. For data to be machine readable, accessible to humans, and usable in different systems with minimal loss of content and functionality, it is essential to include high-quality metadata[35]. The adoption of the Data Catalog Vocabulary[36] metadata standards and mandating the use of open APIs across platforms created by all levels of government will ensure the seamless flow of data and availability of services in all scenarios, including when there is a genuine need for solutions specific to certain locations or jurisdictions[37]. Such efforts will be a more effective use of finite resources as well as expand access to information among the public.

Instituting inclusivity by design

Open data initiatives require an inclusive approach to ensure that the perspectives of marginalised and vulnerable communities are taken into consideration. Given that a great deal of public interest government data is collected from citizens availing social welfare schemes, data in fields like health, education, and employment assistance is likely to primarily reflect people from vulnerable socioeconomic backgrounds.

Comprehensive plans are needed to address data privacy and protection, as well as the underrepresentation and barriers faced by these communities in accessing government information and participating meaningfully in the government's decision-making[38]. These insights should be accounted for in design and user engagement, priority categories in data provisioning, checks and balances for completeness and accuracy of data. Alongside technical aspects like formats or privacy and protection, the language localisation of open data platforms is also essential to ensure a healthy participation by diverse citizens in the creation and use of open data, as recommended by the National Policy on Universal Electronic Accessibility[39].

While the potential of open data to provide substantial public benefit is well established, it is important to be mindful of the power dynamics and contextual nuances of existing social systems to ensure opening up data does not result in harm to vulnerable groups[40]. Moreover, given that public interest data may be skewed in coverage and provide an incomplete picture, it is also critical to be thoughtful in extrapolating to inform large-scale insights [41].

Execute comprehensive personal data protection measures

Beginning in 2017, the Indian government has dedicated concerted effort towards the development of data protection and privacy provisions for individuals. To date, multiple iterations of tentative data legislation, referred to as the Personal Data Protection (PDP) bill, have moved through parliament, presently featuring as a primary agenda item. In the context of the rapid institution of comparable legislation in the European Union, California, and China, as well as the pronounced importance of preserving individual privacy in the COVID-19 era, it is critical that Indian legislators move quickly to pass and enforce a PDP law.

At present, the National Data Sharing and Accountability Policy instructs chief data officers to, at their discretion according to a set of criteria, prepare negative lists of data unfit to share given concerns of personal privacy, confidentiality, and national security. Given longstanding and mounting concerns about data privacy among the bureaucracy and civil society, it is probable that the absence of an extensive personal data protection framework has served as a deterrent for the opening up of public interest data.

Likewise, the passage of PDP will generate explicit regulatory confines and accountability measures that will advance the privacy and security of individual data. Given that a significant portion of open data released by government entities concerns the wellbeing (related to education, livelihood, environment, climate) of the public, it is crucial that such provisions are instituted in a time-sensitive way to enable the protection of marginalised populations while simultaneously unlocking the value of data.

Conclusion

Implementing these measures will expand the scope and strengthen the efficacy of efforts to leverage open data for anticipatory and emergency response in the event of future crises. To this extent, India must not only upgrade data infrastructure to consistently provide real-time, accurate information during an emergency situation, but also ensure sustained funding, interconnectedness of efforts, and inclusive citizen engagement. Such reforms will safeguard the long-term reliability of data and contribute to a societal culture of asking for, providing, reusing, and valuing public data.

From outbreaks of contagious disease to extreme weather, future emergencies are all but inevitable. Nonetheless, it is possible to strengthen Indian society's response—avoiding disaster situations and saving countless lives—by creating and leveraging information systems that are reliable, collaborative, and open.

This piece is based on the discussions of the Open Data Working Group co-hosted by IIC and IDFC Institute.

Endnotes

1. Jhankar Mohta, "While Arvind Kejriwal's 'Corona App' shows ample beds available for Covid-19 patients, hospitals deny," *OpIndia*, 16 April 2021, ,https://www.opindia.com/2021/04/arvind-kejriwal-delhi-corona-app-beds-available-covid-19-patients-hospitals-deny/.

2. Suhasini Raj, "Social Media as 'Godsend': In India, Cries for Help Get Results, *New York Times,* May 3, 2021, https://www.nytimes.com/2021/05/03/world/asia/india-covid-social-media-aid.html.

3. Ministry of Electronics and Information Technology, Government of India, *Implementation Guidelines for National Data Sharing and Accessibility Policy (NDSAP)* (National Informatics Centre: 2015) ahttps://data.gov.in/sites/default/files/NDSAP%20Implementation%20Guidelines%202.4.pdf.

4. Open Government Data Platform India, "Open Government Data (OGD) Platform India," National Informatics Centre (NIC), Ministry of Electronics & Information Technology, Government of India, https://data.gov.in/.

5. Pune DataStore, "PMC Open Data Store" Pune Municipal Corporation, http://opendata.punecorporation.org/Citizen/User.

6. India Observatory, "India Observatory :: Home", Foundation of Ecological Security, https://www.indiaobservatory.org.in/.

7. India Data Portal, "IDP | Home Page," Bharti Institute of Public Policy, Indian School of Business, https://indiadataportal.com.

8. Jessica Alexander, "Then and Now: 25 years of disasters, responses, and risk management," *The New Humanitarian*, April 1 2021, https://www.thenewhumanitarian.org/feature/2021/4/1/25-years-of-disasters-responses-and-risk-management.

9. NYC Open Data, City of New York, https://opendata.cityofnewyork.us/.

10. New York City Open Data Team, Ope*n Data for All 2020 Report | Open Data Connecting New Yorkers*, 2020,https://opendata.cityofnewyork.us/wp-content/uploads/2020/09/2020_OpenDataForAllReport_Full.pdf.

11. PetaBencana.id, https://info.petabencana.id/.

12. Dengue Clusters, National Environment Agency Singapore, https://www.nea.gov.sg/dengue-zika/dengue/dengue-clusters.

13. OD Impact, Singapore's "Dengue Cluster Map," GovLab, https://odimpact.org/case-singapores-dengue-cluster-map.html.

14. Minaketan Behera and Preksha Dassani, "Livelihood Vulnerabilities of Tribals during COVID-19," *Economic and Political Weekly,* March 13, 2021, https://www.epw.in/journal/2021/11/commentary/livelihood-vulnerabilities-tribals-during-covid-19.html ; "Nearly One-Third of U.S. Coronavirus Deaths Are Linked to Nursing Homes", *New York Times,* June 1, 2021, https://www.nytimes.com/interactive/2020/us/coronavirus-nursing-homes.html.

15. Gender equality, "Fewer women than men will regain employment during the COVID-19 recovery says ILO," International Labour Organization, https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_813449?lang=en.

16. "Covid-19 has reversed years of gains in the war on poverty," *The Economist,* September 26, 2020, https://www.economist.com/leaders/2020/09/26/covid-19-has-reversed-years-of-gains-in-the-war-on-poverty.

17. World Wide Web Foundation, *Women's Rights Online Digital Gender Gap Audit Scorecard - India,* 2016, http://webfoundation.org/docs/2016/09/WF_GR_India.pdf.

18. Krish Chetty et al., "Bridging the digital divide: measuring digital literacy" *Economics* 12, no.1 (2018), https://doi.org/10.5018/economics-ejournal.ja.2018-23.

19. Policy on Open Application Programming Interfaces (APIs) for Government of India, "API Setu," Ministry of Electronics & Information Technology, Government of India, https://apisetu.gov.in/document-central/api-policy/index.html.

20. Data Maturity Assessment Framework, "Data Maturity Assessment Framework," Ministry of Housing and Urban Affairs, Government of India, http://dmaf.mohua.gov.in/; Data Governance Quality Index, "Overview: Data Governance Quality Index," Development Monitoring and Evaluation Office (DMEO), NITI Aayog, https://dmeo.gov.in/content/dgqi-overview.

21. Ministry of Electronics and Information Technology, Government of India, *Strategy for National Open Digital Ecosystems (NODE) Consultation Whitepaper,* 2020, https://static.mygov.in/rest/s3fs-public/mygov_158219311451553221.pdf.

22. Simplifying open data licences, "Why do we need to license?," European Data Portal e-learning programme, https://data.europa.eu/elearning/en/module4/#/id/co-01.

23. Government Open Data License – India, "Government Open Data License - India National Data Sharing and Accessibility Policy Government of India," Open Government Data (OGD) Platform India, https://data.gov.in/government-open-data-license-india.

24. Introduction, "Accessing secure research data as an accredited researcher," Office of National Statistics, UK Statistics Authority, https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#introduction.

25. Case Studies, "Accessing secure research data as an accredited researcher," Office of National Statistics, UK Statistics Authority, https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme#case-studies.

26. Ministry of Electronics and Information Technology, Government of India, *Strategy for National Open Digital Ecosystems (NODE) Consultation Whitepaper.*

27. Neeta Verma and M. P. Gupta, "Open government data: beyond policy & portal, a study in Indian context" (paper presented at ICEGOV '13: 7th International Conference on Theory and Practice of Electronic Governance, Seoul, Republic of Korea, October 22-25, 2013).

28. Development Monitoring and Evaluation Office (DMEO), NITI Aayog, *Data and Strategy Unit: Detailed Terms of Reference (ToR)*, 2021, https://dmeo.gov.in/sites/default/files/2021-08/DGQI_2.0_Data_and_Strategy_Unit_ToR.pdf.

29. Regina Mihindukulasuriya, "India has at least 62 apps to deal with Covid, but they all do the same job mostly," *The Print*, June 13 2020, https://theprint.in/india/india-has-at-least-62-apps-to-deal-with-covid-but-they-all-do-the-same-job-mostly/439040/.

30. Mihindukulasuriya, "India has at least 62 apps to deal with Covid, but they all do the same job mostly".

31. CoWIN, "CoWIN Dashboard," Ministry of Health and Family Welfare, Government of India, https://dashboard.cowin.gov.in/.

32. Interoperability Framework for e-Governance, "Technical Standards - IFEG | e-Governance Standards," STQC Directorate, Ministry of Electronics and Information Technology, Government of India, http://egovstandards.gov.in/technical-standards-ifeg.

33. Policy On Open Standards, "Policy On Open Standards | e-Governance Standards," STQC Directorate, Ministry of Electronics and Information Technology, Government of India, http://egovstandards.gov.in/policy-open-standards.

34. Joshua Tauberer, *Open Government Data: The Book* (Second Edition: 2014), Chapter 6, https://opengovdata.io/2014/permanence-trust-provenance/.

35. NISO Press, Understanding Metadata, *National Information Standards Organization*, United States of America, 2004, https://www.lter.uaf.edu/metadata_files/UnderstandingMetadata.pdf.

36. W3C Working Draft 04 May 2021, "Data Catalog Vocabulary (DCAT) - Version 3," World Wide Web Consortium (W3C), https://www.w3.org/TR/vocab-dcat-3/.

37. Ministry of Electronics & Information Technology, Government of India, "API Setu"

38. Feminist and inclusive dialogue, "Canada's 2018-2020 National Action Plan on Open Government," Open Government, Government of Canada, https://open.canada.ca/en/content/canadas-2018-2020-national-action-plan-open-government#toc11.

39. Ministry of Electronics and Information Technology, Government of India, https://pib.gov.in/newsite/printrelease.aspx?relid=99845.

40. Wright Glover, Pranesh Prakash, Sunil Abraham, and Nishant Shah, *Report on Open Government Data in India*, The Centre for Internet and Society, 2011, *https://cis-india.org/openness/publications/ogd-report/at_download/file.*

41. Megan O'Donnell, Mayra Buvinic, Shelby Bourgault and Brian Webster, *The Gendered Dimensions of Social Protection in the COVID-19 Context,* Centre for Global Development, 2021, *https://www.cgdev.org/publication/gendered-dimensions-social-protection-covid-19-context*

Trends in Lunar Exploration: Examining the Governance Challenges

Nivedita Raju

Any discussion on lunar exploration must begin with an acknowledgement of the special status accorded to the Moon under international law. Indeed, the Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies, (Outer Space Treaty) includes a specific section, namely paragraph 2 of Article IV, on lunar security. An entire agreement dedicated to activities on the Moon has also been adopted, with 18 ratifications and four signatories as of 2021. While few states are party to the latter, a distinction can nonetheless be made between space and lunar activities. This manifests prominently in the obligation to use the Moon for 'exclusively' peaceful purposes under the Outer Space Treaty[1]. Article IV additionally prohibits the "establishment of military bases, installations and fortifications, the testing of any type of weapons and the conduct of military manoeuvres on celestial bodies" [2]. This is followed by express permission for "military personnel for scientific research or any other peaceful purposes, and any equipment or facility necessary for peaceful exploration" [3]. This provision recognises that the Moon has a higher threshold of demilitarisation, unlike the rest of outer space.

In this legal backdrop, 2021 has witnessed multiple stakeholders surging ahead with lunar exploration plans. Two parallel international partnerships for lunar exploration have emerged—the US-led Artemis Accords and the International Lunar Research Station (jointly led by China and Russia). These partnerships raise several questions about lunar governance. This essay presents an overview of these growing partnerships and then outlines national lunar ambitions. The paper examines key governance concerns for lunar exploration in light of these initiatives. The paper concludes by proposing solutions that would contribute towards a safe, secure, and sustainable future for lunar stakeholders.

International Lunar Partnerships

Artemis Accords

In 2020, the US's National Aeronautics and Space Administration (NASA) introduced the Artemis Accords, a set of guiding principles for cooperation in civil exploration and use of the Moon and other celestial bodies [4]. As of July 2021, twelve states have signed the accords[5]. Since its adoption, the accords have received considerable scrutiny due to the introduction of new concepts implementing ambiguous provisions under the Outer Space Treaty. This includes Section 11 of the accords pertaining to "deconfliction," which proposes creating safety zones. The accords find the legal basis of this section in Article IX of the Outer Space Treaty regarding the obligation to conduct activities "with due regard to the corresponding interests" of other states and also contains language on "potentially harmful interference"[6]. Since the wording of Article IX itself has been criticised for subjectivity[7], there is no clarity on how to effectively implement safety zones without prejudice to rights of other states.

Furthermore, although the Artemis Accords have no binding effect (being 'principles'), it is imperative to acknowledge that they have the ability to eventually develop into norms of customary international law. It appears that countries that wish to partner with the US in future lunar exploration missions must first sign the accords. States' acceptance of the interpretation of space law as suggested in the accords can subsequently crystallise into custom upon generating sufficient state practice and acceptance of these positions as law. For this reason, the accords are viewed by some as a diplomatic tool of the US to influence international space law[8].

International Lunar Research Station

In March 2021, China and Russia announced a memorandum of understanding to collaborate on the International Lunar Research Station (ILRS)[9]. According to the statement issued by China National Space Administration (CNSA) and Roscosmos State Corporation for Space Activities, the ILRS will be designed for "multi-discipline and multi-purpose scientific research activities, including exploration and use of the Moon, moon-based observation, fundamental research experiments and technology verification, with the capability of long-term unmanned operation with the prospect of subsequent human presence"[10]. The ILRS partnership has reportedly approached the European Space Agency, Thailand, the United Arab Emirates and Saudi Arabia for participation [11].

China and Russia have also presented a roadmap on the phased development of the ILRS. The first stage is reconnaissance, involving the design of the ILRS, selection of sites, and technology verification for high-precision soft-landing[12]. The second phase is construction, which aims to begin in 2026 and includes technology verification for the command centre, lunar sample return, and the start of joint operations[13]. The final stage, utilisation, will commence in 2035, heralding crewed missions to the Moon[14].

The timing of the ILRS announcement indicates that the initiative was created as a response to the US-led Artemis Accords, substantiated by reports that discussions to involve Russia in the accords were unsuccessful[15]. At the same time, the US cannot collaborate with China due to the 'Wolf Amendment', a law that prohibits NASA from partnering with China or Chinese agencies without explicit approval from the US Congress and the Federal Bureau of Investigation[16]. Consequently, in the absence of any dedicated coordination between the ILRS and the Artemis Accords, a toxic rivalry between the two initiatives is likely.

National Lunar Ambitions

Space powers

With several ongoing initiatives, the US has emerged as a leading contender in lunar exploration. The Gateway, a station intended for lunar orbit, is officially described as a "multi-purpose outpost orbiting the Moon that provides essential support for sustainable, long-term human return to the lunar surface and serves as a staging point for deep space exploration"[17]. The Artemis programme will play a significant role in the Gateway: facilitating collaboration with other space agencies. Additionally, in 2020, NASA began soliciting invitations for the collection of lunar regolith and has since awarded contracts worth US$25,001 to four companies—Lunar Outpost (Golden, Colorado), Masten Space Systems (Mojave, California), ispace Europe (Luxembourg), and ispace Japan (Tokyo)[18]. Each firm will collect the lunar regolith and provide evidence of

the collection, following which an "in-place" transfer of ownership to NASA will take place[19]. "After [the] ownership transfer, the collected material becomes the sole property of NASA for the agency's use under the Artemis program"[20]. NASA's solicitation of lunar regolith coupled with the Artemis Accords ostensibly aim to normalise extraction and utilisation of space resources. The legal scope of this issue is currently under debate at the United Nations Committee on the Peaceful Uses of Outer Space (UNCOPUOS).

Meanwhile, China has adopted a phase-wise approach to lunar exploration. Phases I and II of the Lunar Exploration Program involved the launch of lunar orbiters Chang'e 1 and Chang'e 2, soft-landing and deployment of rovers on the lunar surface, and sample-return. Phase III concluded with Chang'e 5's triumphant return in 2020[21]. Phase IV consists of Chang'e 6 to collect samples of lunar rocks from the South Pole-Aitken basin[22]; Chang'e 7, which will conduct examination of sites for resources and study for the lunar base; and Chang'e 8, which is expected to test key technologies that will lay the groundwork for a crewed research base on the moon[23]. The construction of the ILRS will begin at the completion of the fourth phase. This approach indicates that China aims to conduct crewed missions to the Moon in the long run, while the current focus is on the study and extraction of resources.

How will the parallel initiatives on the Moon take place when multiple states are involved? The nature of the lunar environment indicates that activity is likely to be concentrated in certain areas due to the availability of resources and distinct advantages of locations. This will pose a key challenge to governance.

Russia has expressed a keen interest to continue lunar exploration through its Luna programme, a series of robotic spacecraft missions that ended with Luna-24 in 1976[24]. The new Luna-25 mission is intended to launch in May 2022[25] to investigate ice reserves under the lunar surface[26]. Further plans for lunar exploration include Luna-26, Luna-27 and Luna-28, involving investigation of lunar ice and sample-return[27]. A national lunar policy and strategy is reportedly under development and will likely be released soon[28]. Similar to China, its partner in the ILRS, Russia's latest mission focuses on lunar sub-surface resources. Russia also intends to study the hazards of lunar dust in the upcoming mission[29]. Given the focus on lunar resources, it appears that crewed missions are not an immediate priority for Russia.

India will have a significant role to play in international partnerships as a more recent space power with a desire to engage in lunar exploration. After the close miss of Chandrayaan-2 in 2019, India's Chandrayaan-3 mission is now scheduled for 2022[30]. India's space and lunar ambitions hint at the need to rethink both political strategy and domestic policy.

The Indian private sector is brimming with untapped potential, currently unfulfilled due to a lack of domestic legislation. For instance, the proposed SpaceCom policy is aimed at promoting private sector participation, yet there is no clarity on how private entities can partner with the Indian Space Research Organisation (ISRO), nor is there any mechanisms (such as an open bidding process) that will ensure a level playing field. Additionally, there is no incentive for India's private sector in the form of intellectual property for space ventures as ISRO automatically claims ownership of such rights. In the absence of enabling regulations, the Indian private sector will not be sufficiently incentivised to participate.

Rethinking political strategy is also important for India, in light of the Artemis programme and ILRS. Given ongoing tensions with China, it is unlikely that India will join the ILRS. While India and Russia continue to share close relations, particularly in space and defence, Chinese-Russian relations have arguably spurred India into developing space capabilities independently, reflected in the Indian-manufactured orbiter, lander and rover for the Chandrayaan-2 mission. At the same time, India and the US share concerns regarding China's position both globally and in the broader Indo-Pacific region, resulting in close cooperation in the security context. This is evident from the Quadrilateral Security Dialogue and the bilateral defence pact for sharing satellite data. India's position between the ILRS and Artemis initiatives, therefore, poses an interesting example for other states intent on pursuing lunar exploration.

Emerging contenders

The capabilities of other states in the international space sector have grown significantly as well. France, Japan and Germany have reorganised their domestic units with a military focus on space[31]. South Korea and Australia have notable technological capacity and lunar ambitions, evidenced by recent attempts at cooperation on lunar exploration[32].

Recommendations have been made by experts for such states to change the dynamics of the security stalemate by grouping together[33]. This proposal would be effective, and certainly applies to the need for enhanced security and stability in lunar activities. The challenge in the lunar context, however, is that the middle powers, including emerging space contenders, may face the ILRS-Artemis 'divide', as some of these countries have already signed the Artemis Accords[34]. Countries may be forced to align with either initiative on political lines, unless it is clarified how cross-collaboration can be implemented. Additional issues will arise for emerging contenders like Australia, a signatory to the Artemis Accords and a party to the Moon Agreement, which contains provisions on the utilisation of space resources. It is unclear how such states will navigate obligations in a manner consistent with both agreements[35].

Key Governance Concerns for Lunar Exploration

Political fragmentation

States' political commitments to new binding instruments in the space sector have considerably weakened. At this critical juncture, the Artemis Accords and the ILRS partnerships present a political conundrum for all other states with lunar ambitions. Both partnerships present opportunities for collaboration between multiple countries, fostering a culture of cooperation in the global space sector. However, it is important to execute these partnerships in a non-exclusionary and cooperative manner.

A series of questions remain unanswered: Will there be negative consequences for space agencies who join only one of the two lunar partnerships? What if a state does not join either initiative? Will such a state automatically be denied collaboration on lunar exploration with the US or Russia and China? The signatories to the Artemis Accords face these questions. Could there be a contradiction between the interpretations of Outer Space Treaty obligations in the Artemis Accords and the ILRS if countries were to join both initiatives, even as the legality undergoes

debate at multilateral forums? Hypothetically, if a safety zone is created on the lunar surface for mining activities under the Artemis Accords, what happens if a state participating in the ILRS in proximity to this zone refuses to recognise it? A fractured lunar future will only erode predictability and certainty in lunar activities, impacting both safety and security. The two partnerships must, therefore, conduct dialogue to ensure cooperation and compatibility.

Rising mistrust

The risk of fragmentation has the potential to breed further mistrust and suspicion in lunar activities. Regrettably, the last decade reflects the unwillingness of states to commit to new measures in the space-security realm. This is evidenced by the EU-led Code of Conduct, the Russia-China Draft Treaty on Prevention of the Placement of Weapons in Outer Space, and of the Threat or Use of Force Against Outer Space Objects, and the deadlock at the Conference of Disarmament. The UK-led UN General Assembly Resolution 75/36 has reinvigorated interest via a new approach at the multilateral level, using a behaviour-specific lens[36]. Such an initiative is much needed. However, in the interim, there is still an urgent need to introduce transparency while space activities, especially lunar activities, continue to proliferate.

The 1994 and 2013 Groups of Governmental Experts on Transparency and Confidence-building Measures (TCBMs) clearly express that states are hesitant to develop or implement new measures due to the misconception that transparency disadvantages the enacting nation by exposing vulnerabilities[37]. This reasoning is erroneous, as TCBMs are extremely beneficial to both the enacting state and the entire space community as a whole. Currently, with tensions rising between powerful lunar stakeholders, TCBMs can convey the intent behind an activity, thereby reducing suspicions about a rival. New measures must be developed for the short-term, while long-term measures undergo consideration.

Competition in concentrated areas

A challenge specific to governance in lunar exploration is that new initiatives will not be evenly spread out across the Moon and will be clustered in locations that offer the most benefits. For instance, the Peaks of Eternal Light near the lunar poles receive sunlight almost continuously, providing advantages for observational purposes and potential use as a stable solar power source[38]. Additionally, the topography of the Moon is such that physical resources, including thorium, uranium and helium-3 are available in greater amounts at specific locations[39]. In the absence of any guidance on access, benefit-sharing and coordination, these concentrated areas are likely to magnify competition, increase risks of misunderstanding and widen the scope for conflict.

The need for an international regime for resources and benefit-sharing is evident. At the sixtieth session in June 2021, the UNCOPUOS Legal Subcommittee established a working group to examine potential legal models for "activities in exploration, exploitation and utilization of space resources"[40]. At the August 2021 session, a proposal for the mandate, terms of reference and methods of work for the working group under a five-year plan was submitted[41]. The mandate states that the group will study existing legal frameworks, assess the benefits of a new framework, and develop a set of recommended principles for such activities[42]. On the basis of the agreed mandate, the working group will agree on a detailed workplan and methods of work in 2022[43].

This is a positive development from a governance perspective, although it remains to be seen whether states can expeditiously reach consensus on a new regime.

First-mover advantage

The term 'first-mover' often refers to circumstances where developed nations unfairly reap benefits to the detriment of others. However, there is scope for such nations to not only gain tangible benefits when competing in resource-rich areas, but to also play a leading role in shaping norms surrounding lunar activities. The legal ambiguities under the space treaties hint at the power of custom in international law. Custom, a source of international law, refers to the general practices of states accepted as law [44]. The two elements of custom are comprised of uniform and consistent state practice, with the recognition that such practice is observed by states as a legal obligation[45]. Consequently, 'first-mover' states have the advantage of defining how the space treaties should be implemented. Such power weighs heavily in favour of countries with sophisticated capabilities and lunar ambitions. Other states, particularly those from the Global South, have little scope to participate in shaping lunar norms, especially if this occurs through state practice in the ILRS and Artemis partnerships. This will subsequently broaden the gap between the Global North and Global South at the international level. Both lunar partnerships must actively include developing nations and represent their perspectives to address the power imbalance associated with this first-mover advantage. This will prevent developing countries from being excluded from shaping lunar policy.

Diversified activities and stakeholders

Foreseeable activities on the Moon are wide-ranging, from resource utilisation to habitable lunar bases. Each of these activities reveal varying interests, which are unique to the activity in question. Furthermore, these interests come from different types of stakeholders. Non-state actors, such as companies and civil society organisations, have begun to play an active role in the space sector, particularly lunar activities. The hurdle in this case is creating avenues for expression and policy engagement from such stakeholders. This is a challenge that urgently requires attention because stakeholders from the private sector can be extremely valuable in creating norms for responsible behaviour, for example, through the implementation of best practices, industry standards and voicing perspectives of underrepresented groups. The potential of the private sector in this regard is evidenced by several initiatives, such as The Hague International Space Resources Governance Working Group and For All Moonkind, which are attempting to shape the discussion on different topics[46]. The creation of the Breaking Ground trust is another example of building policy for the equitable use of lunar resources pursuant to NASA's regolith solicitation[47]. The UN General Assembly Resolution 75/36 had earlier invited inputs from the private sector, which facilitated an expression of views from non-governmental organisations, including think tanks and international entities like the International Committee of the Red Cross and the United Nations Institute for Disarmament Research. Yet, there continues to be little scope for non-state actors to play an active role in existing state-oriented institutional design. New platforms should be considered to enable private sector stakeholders to join discussions on policy solutions for lunar governance challenges.

Solutions

Clarifying the lex lata governing the Moon

A comprehensive review of the applicable normative framework for the Moon is lacking. Paragraph 2 of Article IV of the Outer Space Treaty establishes the Moon's special status. Certainly, the meaning of the expression 'peaceful purposes' has undergone extensive discussion and has an expansive interpretation today. However, there is scant literature that examines precisely how countries have responded to the obligation of 'exclusively' peaceful purposes. In the absence of this understanding, several questions emerge. For instance, how does 'exclusively' peaceful extend to the space activities in lunar orbits? Would 'exclusively' peaceful subsequently extend to any lunar resources extracted from the Moon, as such materials are part of the lunar surface? Furthermore, considering potential norm-creation in upcoming lunar activities, how can the conduct of non-participating countries be included? These questions require urgent clarification.

The Manual on International Law Applicable to Military Uses of Outer Space is an ongoing initiative that presents the *lex lata* for outer space, including the Moon. In addition, an inquiry dedicated to the normative framework of the Moon will clarify the precise extent of demilitarisation and the legal scope of future missions. This will more effectively support the design and implementation of policies for both international and national lunar exploration.

Lunar-specific TCBMs

Binding legal regimes to tackle challenges related to lunar exploration are the preferred outcome. However, in the current political climate, where 'hard' binding measures continue to be controversial, short-term measures to rebuild trust and facilitate communication are vital in the interim. These measures should be given adequate policy focus to prevent existing tensions from being escalated by an actor based on the limited information of another actor's conduct. Developing TCBMs specifically for lunar activities can ensure the communication of intent to allay suspicions. These measures could manifest in multiple ways and be undertaken unilaterally, or on a bilateral and multilateral basis. One example would be to improve registration practices. Weak compliance with registration under the Registration Convention is amplified in the lunar context, as there is no mention of specific obligations to register interactions between objects on the lunar surface or in cis-lunar space[48]. TCBMs on the registration of lunar activities can thus be a useful theme in which to invest political and technical effort.

The sharing of space situational awareness (SSA) data is another critical area that requires focus. While some bilateral and multilateral efforts have been initiated, a dedicated effort towards the collection, processing and verification of SSA data between different lunar stakeholders has the dual benefit of lowering the risk of behaviour being misconstrued and increasing accountability for the stakeholders involved[49].

Cooperation protocols between Artemis Accords and ILRS

States leading the international partnerships for lunar explorations must prioritise the development of protocols for cooperation and information sharing. Cooperation between rival

countries is not impossible, and has been an integral component of past space operations. Indeed, precedent can be derived from the cooperation mechanisms present in the International Space Station. While there is a great deal of pride and patriotism associated with lunar missions, cross-partnership cooperation between states will prove highly constructive. Past instances of cooperation are reflected in China's Chang'e 4 landing, when NASA coordinated with CNSA to provide monitoring and observational support for the mission[50].

Cooperation protocols can include fundamental "ground rules" for all parties in the partnerships, perhaps beginning with humanitarian protections. Current provisions in the Outer Space Treaty and the Rescue and Return Agreement provide certain obligations for states to assist astronauts from other states[51]. Yet, a distinction between astronauts or personnel of a spacecraft and spaceflight participants persists, especially on the brink of a budding space tourism industry[52]. ILRS-Artemis cooperation can commence discussion on this point, where parties agree to provide mutual aid in case of distress faced by participants in either initiative. Similarly, there can be communication protocols for the periodic exchange of information, in addition to reinforcement of the obligation regarding visitation rights of installations, equipment and stations under the Outer Space Treaty[53].

Conclusion

This paper presented trends in lunar exploration through the lens of international partnerships amid growing national ambitions. Evidently, the rate of innovation has far outpaced multilateral policymaking machinery, given the lack of regulations governing these activities. New solutions to advance regulation in this context must account for political realities. For this reason, lunar specific TCBMs are proposed as a starting point. The reinforcement of trust and confidence coupled with an ILRS-Artemis protocol for cooperation will ensure that the two partnerships can be conducted simultaneously, even in the same resource-rich and advantageous lunar areas. These solutions are aimed at reducing the scope for toxic rivalry to escalate, and must get immediate attention. These measures should additionally be accompanied by a comprehensive inquiry into the normative framework of the Moon, as clarifying the *lex lata* will guide the design and implementation of policies for future activities. To ensure that the Moon continues to enjoy its unique status, this exercise can begin with a study on state practice on the obligation of use for 'exclusively' peaceful purposes.

Parallel to these short-term measures, efforts must continue multilaterally at the UN level. Developing measures under UN General Assembly Resolution 75/36 in conjunction with continued discussions at the UNCOPUOS will be essential. The dedicated pursuit of both short- and long-term measures will ensure safe, secure, and sustainable lunar exploration.

Endnotes

1. Article IV, *Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, Including the Moon and Other Celestial Bodies*, 27 January 1967, 610 UNTS 205 [*Outer Space Treaty*].
2. Art. IV, *Outer Space Treaty*.
3. Art. IV, *Outer Space Treaty*.

4. NASA, Artemis Accords: Principles for Cooperation in the Civil Exploration and use of the Moon, Mars, Comets and Asteroids for Peaceful Purposes, 2020 https://www.nasa.gov/specials/artemis-accords/img/Artemis-Accords-signed-13Oct2020.pdf

5. NASA, *Artemis Accords*, 2020 https://www.nasa.gov/specials/artemis-accords/index.html

6. Art. IX, *Outer Space Treaty*.

7. Nivedita Raju, 'A Proposal for a Ban on Destructive ASAT Testing: A Role for the European Union?' *EU Non-Proliferation and Disarmament Consortium Papers*, no. 74, 2021.

8. European Space Policy Institute (ESPI), *Artemis Accords: What implications for Europe?*, ESPI Brief no. 46, 2020, https://espi.or.at/downloads/send/5-espi-executive-briefs/554-artemis-accords-what-implications-for-europe; European Parliament, Policy Department for External Relations, *The European space sector as an enabler of EU strategic autonomy*, 2020, https://www.europarl.europa.eu/RegData/etudes/IDAN/2020/653620/EXPO_IDA(2020)653620_EN.pdf

9. Andrew Jones, "China, Russia enter MoU on international lunar research station," *SpaceNews*, March 9, 2021, https://spacenews.com/china-russia-enter-mou-on-international-lunar-research-station

10. China National Space Administration (CNSA), Government of the People's Republic of China, http://www.cnsa.gov.cn/english/n6465652/n6465653/c6811967/content.html.

11. Tereza Pultanova, "Russia, China reveal moon base roadmap but no plans for astronaut trips yet" *Space*, June 17 2021 https://www.space.com/china-russia-international-lunar-research-station

12. China National Space Administration (CNSA), Government of the People's Republic of China, http://www.cnsa.gov.cn/english/n6465652/n6465653/c6812150/content.html

13. CNSA, "International Lunar Research Station (ILRS) Guide for Partnership"

14. CNSA, "International Lunar Research Station (ILRS) Guide for Partnership"

15. Moon Dialogs, *Peaceful Moon: International Collaboration for Lunar Bases*, 2021 https://www.moondialogs.org/events/moon-dialogs-11-peaceful-moon-salon-fractured-lunar-futures.

16. Makena Young, "Bad Idea: The Wolf Amendment (Limiting Collaboration with China in Space," *Defense360*, December 4, 2019, https://defense360.csis.org/bad-idea-the-wolf-amendment-limiting-collaboration-with-china-in-space/

17. NASA, "Gateway" NASA https://www.nasa.gov/gateway

18. NASA, Government of the United States of America, https://www.nasa.gov/press-release/nasa-selects-companies-to-collect-lunar-resources-for-artemis-demonstrations.

19. NASA, "NASA Selects Companies to Collect Lunar Resources for Artemis Demonstrations",

20. NASA, "NASA Selects Companies to Collect Lunar Resources for Artemis Demonstrations",

21. "China's Chang'e 5 mission a success, spacecraft brings home first moon samples in 40 years," *Economic Times*, December 17, 2020, https://economictimes.indiatimes.com/news/international/world-news/chinas-change-5-mission-a-success-spacecraft-brings-home-first-moon-samples-in-40-years/articleshow/79772341.cms?from=mdr.

22. Andrew Jones, "China unveils ambitious moon mission plans for 2024 and beyond," *Space.com*, October 19, 2020, https://www.space.com/china-planning-future-moon-missions-change-7

23. Jones, "China unveils ambitious moon mission plans for 2024 and beyond"

24. Jones, "China unveils ambitious moon mission plans for 2024 and beyond"

25. "Launch of Russia's Luna-25 automatic station to Moon rescheduled for 2022" *Russian News Agency TASS,* August 20, 2021, https://tass.com/science/1328449.

26. Meghan Bartels, "Russia is going back to the moon this year," *Space*, April 15, 2021, https://www.space.com/russia-luna-25-returning-to-moon.

27. Bartels, "Russia is going back to the moon this year"

28. "*Peaceful Moon: International Collaboration for Lunar Bases*,"

29. IKI, Russian Space Research Institute, "Russian Moon Exploration Program," Russian Academy of Science, http://www.iki.rssi.ru/eng/moon.htm

30. "Isro's Chandrayaan-3 launch next year: Govt," *The Hindustan Times,* March 25. 2021 https://www.hindustantimes.com/india-news/isros-chandrayaan-3-launch-next-year-govt-101616619292766.html

31. Sabine Siebold, "New German space command to tackle Russian, Chinese threat, overcrowding" *Reuters* July 13, 2021, https://www.reuters.com/business/aerospace-defense/new-german-space-command-tackle-russian-chinese-threat-overcrowding-2021-07-13/

32. For example, see NASA's support for upcoming robotic lunar exploration mission of South Korea; "NASA Selects Nine Scientists to Join Korea Pathfinder Lunar Orbiter Mission," https://www.nasa.gov/feature/nasa-selects-nine-scientists-to-join-korea-pathfinder-lunar-orbiter-mission

33. Rajeswari Pillai Rajagopalan, "Changing Space Security Dynamics and Governance Debates," in *Commercial and Military Uses of Space*, ed. Melissa de Zwart and Stacey Henderson (Adelaide: Springer Publishing, 2021) 165.

34. NASA, "Artemis Accords"

35. Nivedita Raju, Heloise Vertadier, "The Role of Customary International Law in Contemporary Space Activities," *Journal of Space Law*, 44, no. 3 (2021) (*forthcoming*).

36. UN General Assembly Resolution, 'Reducing space threats through norms, rules and principles of responsible behaviours," UN Doc A/RES/75/36 (2020).

37. UN General Assembly Resolution, "Prevention of arms race in outer space, Study on the application of confidence-building measures in outer space," UN Doc A/48/305 (15 October 1993) at para 305

38. Martin Elvis et al, "Concentrated Lunar Resources: Imminent Implications for Governance and Justice," *Philosophical Transactions of The Royal Society: A Mathematical Physical and Engineering Sciences*, 379 (2021).

39. Elvis et al "Concentrated Lunar Resources: Imminent Implications for Governance and Justice".

40. UNCOPUOS, "General exchange of views on potential legal models for activities in exploration, exploitation and utilization of space resources," 60th Sess, UN Doc A/AC.105/C.2/L.314/Add.8 (2021).

41. UNCOPUOS, 'Proposal on the mandate, terms of reference, and workplan and methods of work for the working group established under the Legal Subcommittee agenda item entitled "General exchange of views on potential legal models for activities in the exploration, exploitation, and utilization of space resources"', 60th Sess, UN Doc A/AC.105/2021/CRP.11 (2021).

42. UNCOPUOS, 'Proposal on the mandate, terms of reference, and workplan and methods of work for the working group established under the Legal Subcommittee agenda item entitled "General exchange of views on potential legal models for activities in the exploration, exploitation, and utilization of space resources"'

43. UNCOPUOS, "Proposal on the mandate, terms of reference, and workplan and methods of work for the working group established under the Legal Subcommittee agenda item entitled "General exchange of views on potential legal models for activities in the exploration, exploitation, and utilization of space resources"

44. Article 38(b), *Statute of the International Court of Justice*, 26 June, 1945 (entered into force 24 October 1945).

45. Raju "The Role of Customary International Law in Contemporary Space Activities,".

46. The Hague International Space Resources Governance Working Group, https://www.universiteitleiden.nl/en/law/institute-of-public-law/institute-of-air-space-law/the-hague-space-resources-governance-working-group; For All Moonkind, https://www.forallmoonkind.org/

47. Breaking Ground, https://breakingground.space/.

48. Nivedita Raju, "Transparency and Confidence-Building Measures for Lunar Security," *Open Lunar Foundation*, May 19, 2021.

49. Raju "Transparency and Confidence-Building Measures for Lunar Security,"

50. "NASA's lunar orbiter has its third, overhead look on China's Chang'e-4 probe" Xinhua February 16, 2021, http://www.xinhuanet.com/english/2019-02/16/c_137825763.htm

51. Art. V, *Outer Space Treaty*; Arts. 1-4, *Agreement on the Rescue of Astronauts, the Return of Astronauts and the Return of Objects Launched Into Outer Space*, 22 April, 1968, (entered into force 3 December 1968).

52. See latest FAA order which prescribes eligibility requirements for Commercial Astronaut Wings, including a demonstration of "activities during flight that were essential to public safety, or contributed to human space flight safety," US Federal Aviation Administration, *Order 8800.2, FAA Commercial Space Astronaut Wings Program,* 20 July, 2021.

53. See Art. XII, *Outer Space Treaty*, which states, "[a]ll stations, installations, equipment and space vehicles on the Moon and other celestial bodies shall be open to representatives of other States Parties to the Treaty on a basis of reciprocity." This has been recommended as "familiarization visits" under the latest report of the UNGA Resolution 75/36.

About the Editors & Authors

Trisha Ray

Trisha Ray is an Associate Fellow at the Centre for Security, Strategy and Technology, Observer Research Foundation, where she works at the intersection of emerging technologies and geopolitics. She is also Chair of the 2021 edition of CyFy, ORF's flagship technology conference.

Rajeswari Pillai Rajagopalan

Rajeswari Pillai Rajagopalan is the Director of the Centre for Security, Strategy and Technology (CSST) at the Observer Research Foundation, New Delhi. Dr. Rajagopalan joined ORF after a five-year stint at the National Security Council Secretariat (2003-2007), Government of India, where she was an Assistant Director.

Philip Reiner

Philip Reiner is the Chief Executive Officer of the Institute for Security and Technology, a global non-profit that serves as the bridge between technologists and national security policy makers to fix tech-driven emerging security threats.

Abhinav Verma

Abhinav Verma is a lawyer specialising in international law, and a public policy professional working at the intersection of social innovation, strategic social investment, and technology for good.

Arindrajit Basu

Arindrajit Basu is Research Lead at the Centre for Internet and Society, and his writing covers geopolitics, constitutional law, and technology.

Victoria Samson

Victoria Samson is the Washington Office Director for Secure World Foundation and has twenty years of experience in military space and security issues.

Abagail Lawson

Abagail Lawson is an Associate Fellow and Program Manager at ORF America, and her writing covers issues at the intersection of technology and policy.

Smriti Parsheera

Smriti Parsheera is a Fellow with the CyberBRICS Project hosted by FGV Law School, Brazil and a PhD candidate at the Indian Institute of Technology Delhi.

James A. Lewis

James A. Lewis is the Senior Vice President and the Director of the Technology and Public Policy Program at the Center for Strategic and International Studies.

Nikhila Natarajan

Nikhila Natarajan is Senior Programme Manager for Media and Digital Content with ORF America.

Paul Cadario

Paul Cadario is the University of Toronto's Distinguished Fellow in Global Innovation at the Munk School of Global Affairs and Public Policy and the Faculty of Applied Science & Engineering.

Daniel Porras

Daniel Porras is Director of Strategic Partnerships and Communications at the Secure World Foundation. He works to develop partnerships and conduct outreach for the development of sustainable norms of behaviour in space. He is also a Non-Resident Fellow at the UN Institute for Disarmament Research, where he focuses on political and legal issues surrounding space security.

Nisha Holla

Nisha Holla is Visiting Fellow at the Observer Research Foundation and Technology Fellow at the Center for Cellular and Molecular Platforms, Bengaluru.

Sam Naufeld

Sam Naufeld is currently working at SOSV, a global venture capital firm that operates early-stage startup development programmes. He was previously a project manager at the Open Data Working Group co-hosted by the International Innovation Corp.

Sridhar Ganapathy

Sridhar Ganapathy is a Senior Associate at IDFC Institute who works on data science.

Nivedita Raju

Nivedita Raju is a researcher at Stockholm International Peace Research Institute (SIPRI) with a focus on space security and disarmament education.