Decoding Digital Authoritarianism

Report prepared for Global Affairs Canada*

March 2023

Berggruen Institute | Canada | UCLA Institute for Technology, Law & Policy | USC Dornsife Center on Science, Technology, and Public Life

*Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

Table of Contents

Part III Global Digital Finance and Democratic In/Exclusion

Part IV New Tech: Despotic Blockchains and Exploitative AI

Executive Summary

Ziyaad Bhorat
Berggruen Institute, University of Southern California, and Harvard University*

Problem

Digital technologies hold tremendous potential to enhance societal wellbeing, including the protection and promotion of democracy and human rights. However, citizens, political leaders, and states are increasingly concerned about the way in which digital technologies can be used to curtail freedoms and oppress foreign and domestic population groups in liberal democratic and autocratic states alike. Digital authoritarianism has emerged as a policy-relevant description of this concern, but suffers from a lack of i) conceptual clarity, ii) relevant stakeholder consensus on areas of focus, and iii) strategic policy guidance for effective response.

Background

More than a third of the world's population now lives under some form of authoritarian rule, and if we look at empirical measures like the Democracy Index, a majority of countries registered a deterioration in their average score or stagnated in 2022 – continuing a downward trend, accelerated by the global pandemic, to the worst average global score since the Index's inception in 2006.[1] At the same time, almost two thirds of the global population, 5.16 billion people, are now connected to the internet.[2] And in 2022, global investment in Artificial Intelligence (AI) stood at $136.55 billion, with growth expected to accelerate further as recent trends with generative AI take shape.[3] Moreover, some social media platforms, like Meta, yield revenues in excess of entire nation states. It is clear therefore that digital tools and platforms have expanded into virtually every area of human life, corresponding to an enormous potential for political, economic, and social discrimination and exclusion. In short, human freedoms are under threat, while new technologies continue a breakneck onward march.

Key political leaders and actors have therefore sounded the alarm. US Secretary of State Anthony Blinken, for example, has called for a response to "abusive technology, including digital authoritarianism" used "to stifle dissent, to surveil and censor".[4] And in 2023, President Joe Biden's White House released its National Cybersecurity Strategy, accusing China of "exporting its vision of digital authoritarianism, striving to shape the global internet in its image and imperiling human rights beyond its borders".[5]

The term "digital authoritarianism" is indebted to earlier ideas of 'networked authoritarianism', which describes an already "authoritarian regime embrac[ing] and adjust[ing] to the inevitable changes brought by digital communications".[6] Digital authoritarianism has therefore become a general term levelled at states like China, Iran, India, and Russia to mean "the use of digital information technology by authoritarian regimes to surveil, repress, and manipulate domestic and foreign populations".[7] Recent discussion has focused on how tools and technologies that include "censorship and automated surveillance systems",[8] present "overlapping and expanding challenges (1) within autocracies, (2) as tools to undermine adversaries, (3) via export to like-minded regimes, and (4) within and by democracies themselves".[9] Scholarship in turn has largely focused on Chinese and Russian models of digital authoritarianism,[10] and digital authoritarian regional practices in the Middle East.[11]

Report Structure

Taking the traditional understanding of digital authoritarianism as a point of departure, this report tasked 9 experts from a variety of regional focus areas and academic disciplines – inter alia comparative and international law, political economy, data studies, and media & communications – with offering their views on what digital authoritarianism means, what empirical factors drive it, and how we can develop theory around it that allows for a more global, inclusive conversation to address the various ways it might manifest in societies. This approach was specifically designed to include novel angles into and/or underrepresented voices in conversations about digital technology and policy, and so these contributions range from the Global South implications of AI 'epistemic authoritarianism', to the urgent (un)democratic potential of central bank digital currencies in global finance.

An overview of the report structure and contributor papers appears below:

Part I: Digital Geopolitics

"Whose Internet? Authoritarianism and the Struggle Over Governance"

Kal Raustiala (UCLA)

A comparative international law perspective on how California versus China models of internet governance emerged and are playing out in differing geopolitical strategic ambitions in multistakeholderism and multilateralism respectively.

"Foreign Actors and Digital Authoritarianism in Africa: Recent Trends on Methods and Their Human Rights Impacts"

Tomiwa Ilori (University of Pretoria)

Shows how foreign actors, including private firms, have sought to export authoritarian methods into Africa along two major trends: networked surveillance and information disorder. Notably, not only China but also Western states play host to these actors.

Part II: Infrastructures of Oppression and Resistance

"An Infrastructural Approach to Digital Authoritarianism"

Mike Ananny (USC)

Replaces the traditional definition's emphasis on tools and technologies with an understanding of digital authoritarianism as the unchecked creation and deployment of oppressive communication infrastructure. Thus, developing effective resistance strategies means acknowledging how these neutral-seeming infrastructure's genealogies—layers of history, skill, judgment, assumption, technique, language, revision, and contestation—are actually deeply political.

"Panopticons and Closing Civic Space: The Building Blocks of Digital Authoritarianism"

Courtney Radsch (UCLA)

Presents the risk complex digital infrastructures pose to democratic values, arguing extraordinary advances in computing power, and 'datafication' have led to:
i) A misalignment of political and economic values perverting the public sphere through disinformation, propaganda, and harassment;
ii) New digital national ID cards, social credit systems, and other forms of surveillance that close civic space, increase human rights abuses, and make anonymity impossible.

Part II: Infrastructures of Oppression and Resistance (continued)

"Protecting Digital Infrastructures, Assets, & Users from Bad Actors: An Empowering Story of JamiiForums from Tanzania"

Patricia Boshe (University of Passau)

A Tanzanian case study on how a democratic state can adopt authoritarian data practices, and yet be successfully resisted by local digital platforms whose data policies and offshore servers are better able to protect users from state overreach.

Part III: Global Digital Finance and Democratic In/Exclusion

"Digital Finance and the Specter of Digital Authoritarianism"

Saule Omarova (Cornell)

Argues for the democratic adoption of a central bank digital currency (CBDC) as failing to do so leaves the global financial system vulnerable to authoritarian states like China who are already investing in this area. Leaving digital currency in the hands of private sector stablecoins moreover presents an internal challenge to democratic sovereignty.

"Undoing Democratic Social Citizenship? The Digitalization of G2P Payments and the Making of Private Digital Authoritarianisms"

Nicholas Bernards (University of Warwick)

Warns that the digitization of social transfers under the auspices of financial inclusion, encouraged by IGOs like the World Bank and large philanthropic nonprofits, can actually remove social protection systems from democratic accountability in favour of control by large private sector monopolies.

Part IV: New Tech: Despotic Blockchains and Exploitative AI

"Technocratic Despotism in The Network State"

Morshed Mannan (EUI)

Examines and critiques the recent libertarian push for a blockchain-enabled 'network state', and how its emphasis on exiting the nation-state fails to account for principles of good political governance and the need for multi-stakeholder perspectives, including the voice of feminist groups and indigenous communities.

"Platform Authority and Data Quality: Who Decides What Counts in Data Production for Artificial Intelligence?"

Julian Posada (Yale)

Argues for a novel conception of 'epistemic authoritarianism' in global digital platforms which rely on exploited labor, especially in Africa and Latin America, to assert 'ground truth' conditions in the generation, annotation, and verification of data meant for AI systems.

General Findings and Recommendations

1. The Multilateral Challenge: How to Keep Friends and Influence People

Experts argue that states like China are increasingly exerting influence in multilateral settings to shape new models of digital governance, while Western liberal democracies are falling back on broader and decentralized multistakeholder approaches.

'Digital authoritarianism' is already geopolitically coded against illiberal non-Western states like China, Russia, India, and Iran, which threatens its usefulness in multilateral initiatives. This is an acute concern moreover for winning over non-aligned or otherwise borderline states, whose reaction to the term might place common anxieties about digital technology second to geopolitical coding.

Recommendation:
States caught between the Global North language of digital authoritarianism and adverse reactions to it in the Global South should coordinate security policy needs and communication with broader diplomatic policy ambitions – especially in multilateral settings. Outside the G7 and groups like the Freedom Online Coalition, the language of digital authoritarianism is likely to frustrate cooperation.

2. Knots of Repression Beyond Authoritarian States

Experts highlighted that what is traditionally understood by digital authoritarianism can manifest in otherwise democratic contexts. The case of Tanzania, for example, shows that repressive digital ***practices*** can occur in democratic states, while the work on digital stablecoins, digital social transfer firms, and global AI platform labor supply, suggest that ***actors*** like private sector monopolies can undermine democracy both domestically and globally.

Digital technologies also present ***new entry points for repression*** that have otherwise been overlooked, such as FinTech privatization of crucial social transfer systems in Africa, effectively removing these from democratic accountability in favor of 'financial inclusion'. Private international philanthropy groups and IGOs like the World Bank can (unwittingly) be complicit actors in these processes.

Democratic states can also host private sector actors that export repressive digital technologies to other states, whether democratic or autocratic. The concentration of private surveillance firms in the EU and US, for example, can lead to an ***export of digital repression originating in democratic states.***

The traditional definition of digital authoritarianism can therefore lead to dangerous blind spots in liberal democratic states as regards digitally-enabled reversals and exports toward authoritarianism.

Recommendation:
Focus on a broader set of non-state actors – especially in the private sector but also N/IGOs– along with practices involved in digital repression, while coordinating on security and (anti-monopoly) industrial policy to avoid centralized 'knots' of power in the digital space.

3. Infrastructures, Not Tools

Experts worry that a definitional and/or focus on specific *tools* and or *technologies* of surveillance, repression, and manipulation can redirect away from broader underlying ***infrastructures*** that enable their emergence. Infrastructures include ***technical*** dimensions that underlie digital technologies, for example server systems; as well as attendant ***social*** dimensions that include "genealogies—layers of history, skill, judgment, assumption, technique, language, revision, and contestation—that may *seem* boring and neutral, but are actually deeply political".

On one hand, an emphasis on technical infrastructures means that digital policymaking ought to coordinate with both security and traditional ICT sector policy. For example, the Tanzanian case study shows how offshore servers used by a local social media platform allowed it to successfully resist repressive state practices. There is strategic value therefore along the technical infrastructure pipeline, which (foreign) actors may be able to exploit according to differing ambitions.

One the other hand, social cultures of digital innovation are also part of infrastructural analysis, and present a need for policy-driven incubation away from the development of "charismatic geniuses and [often libertarian] techno-utopian cultures" that have resulted in single-authority figures in digital technology, towards cultures that emphasize multistakeholder perspectives that include marginalized voices (global/domestic feminist, indigenous groups).

Recommendation:
Increasing coordinated and strategic investments in both global/domestic technical and social digital infrastructure spaces to align values, protect allies, and incubate bottom-up multistakeholder cultures of innovation.

In sum

The traditional conception of digital authoritarianism can occlude broader challenges with digital tools, systems, and infrastructures that present globally with/to a variety of actors that include, but are not exclusively authoritarian nation states.

Opportunity

Redefining and redeveloping the problem of digital authoritarianism in terms of a strategic digital governance framework that focuses on defined outcomes and measures (e.g. levels of digital/tech industry consolidation) with an increased potential for multilateral agreement. Initiatives like the Global Digital Compact can inform this framework, but further work needs to be done to tailor and operationalize digital strategy in domestic and foreign policy decision-making and standards development.

Principal Investigators

Dr. Ziyaad Bhorat is a South African political theorist who works on automated technologies, global digital governance, and democratic politics. His work has been published in academic and general audience outlets across the US, UK, and South Africa, on topics ranging from Aristotle to AI judges. He holds a Ph.D. in Political Science from the University of California, Los Angeles (UCLA), where he also received his M.A. He also holds an M.B.A., and an M.Sc. (African Studies) from the University of Oxford, studying there as a Rhodes Scholar.

Ziyaad is currently a fellow at the Berggruen Institute, USC's Center on Science, Technology, and Public Life, as well as a Technology and Human Rights Fellow at the Carr Center for Human Rights Policy at Harvard University. He has also worked in the digital media and telecoms industries across sub-Saharan Africa and on projects in these regions for companies like The Walt Disney Co.

Martin Rauchbauer works at the intersection of diplomacy, policy, the environment, and the arts. He is the Executive Director of the Djerassi Resident Artists Program and the Co-Founder of the Tech Diplomacy Network. Previously, after serving for two years as Austria's first Tech Ambassador in Silicon Valley. During more than five years as Head of Open Austria and Austrian Consul in San Francisco, he shaped the emerging field of tech diplomacy, engaged in transatlantic digital diplomacy and digital human rights. He also developed digital humanism as a strategic focus of Austrian foreign policy. Martin initiated Open Austria's "Art + Tech Lab", and co-founded the European art + tech + policy initiative "The Grid". he is currently a fellow at the Berggruen Institute focusing his research on how tech governance and diplomacy are conceptually based on our understanding of the human, nature, and technology.

Expert Contributors

Kal Raustiala is the Promise Institute Distinguished Professor of Comparative and International Law at UCLA Law School and Director of the UCLA Ronald W. Burkle Center for International Relations. His research focuses on international law, international relations, and intellectual property. His recent publications include "The Fight Against China's Bribe Machine," *Foreign Affairs*, October 2021 (with Nicolas Barile); "NGOs in International Treatymaking," in *The Oxford Guide to Treaties* (Oxford 2020); " Hollywood is Running Out of Villains," *Foreign Affairs*, August 2020; "Innovation in the Information Age: The United States, China, and the Struggle Over Intellectual Property in the 21st Century," *Columbia Journal of Transnational Law* (June 2020); and "The Second Digital Disruption: Streaming and the Dawn of Data-Driven Creativity," *NYU Law Review* (2019, (with Christopher Sprigman). His books include *The Absolutely Indispensable Man: Ralph Bunche, the United Nations, and the Fight to End Empire* (Oxford 2022); *Global Governance in a World of Change* (Barnett, Pevehouse, and Raustiala, eds, Cambridge, 2021); and *Does the Constitution Follow the Flag? The Evolution of Territoriality in American Law* (Oxford, 2009). *The Knockoff Economy: How Imitation Sparks Innovation* (Oxford, 2012) (with Christopher Sprigman), was translated into Chinese, Korean, and Japanese. In 2016 Professor Raustiala was elected Vice President of the American Society of International Law. He has taught at Yale Law School, Harvard Law School, Columbia Law School, Princeton University, the University of Chicago Law School, Melbourne University, Hebrew University, and the National University of Singapore. A graduate of Duke University, Professor Raustiala holds a J.D. from Harvard and Ph.D. from UC San Diego. Prior to coming to UCLA, he was a research fellow in the Foreign Policy Studies Program at the Brookings Institution and an assistant professor of politics at Brandeis. A life member of the Council on Foreign Relations, he has served on the editorial boards of *International Organization* and the *American Journal of International Law*.

Julian Posada is a Postdoctoral Associate and incoming Assistant Professor of American Studies at Yale University, where he is also a fellow of the Yale Law School's Information Society Project. His research integrates theories and methods from information studies, sociology, and human–computer interaction to study technology and society. He is currently researching the relationship between human labor and data production in the artificial intelligence industry. This project centers on the experiences of outsourced workers in Latin America employed by digital platforms to produce machine learning data and verify algorithmic outputs. Julian's research has been published in several influential journals, including *Information, Communication & Society*, the *Proceedings of the ACM on Human-Computer Interaction*, and in book chapters published by Oxford University Press and SAGE. He is committed to public engagement and has published in *Logic Magazine*, *Bot Populi*, and *Ethics in Context*.

Tomiwa Ilori is a Postdoctoral Research Fellow at the Centre for Human Rights, Faculty of Law, University of Pretoria. His current research focuses on platform governance and accountability, data protection and communication surveillance law. He has worked in various capacities on digital rights and policy-related research and projects including those that focus on the international human rights system and new technologies.

Nick Bernards is an Associate Professor in Global Sustainable Development at the University of Warwick. He is a political economist with research interests in the past and present intersections of labour, finance, and global governance. His work is historically-oriented, with an emphasis on how long-run legacies of colonialism have shaped the present context of sustainable development practice. Prior to starting at Warwick in 2017, he held a Postdoctoral Fellowship in the Department of Political Studies at Queen's University, Canada, funded by the Social Science and Humanities Research Council of Canada. He completed a PhD in International Relations at McMaster University in 2016. Nick has published on a range of issues around labour, finance, and governance including colonial histories, agrarian finance, informal economies, technological change, and international labour regulation. His first book, *The Global Governance of Precarity* (Routledge, 2018), examines the governance of irregular forms of labour in sub-Saharan Africa through a historical study of the activities of the International Labour Organization. His most recent book, *A Critical History of Poverty Finance* (Pluto Press, 2022) looks at the global history of efforts to extend financial services to the poorest. The book puts recent initiatives promoting the use of new financial technologies in the context of a longer history dating back to inter-war colonialism. The book draws on this history as a way of examining the limits of neoliberal models of development, showing how efforts to resolve poverty through the construction of new markets have often exacerbated existing patterns of uneven development.

Dr. Morshed Mannan is a Research Fellow at the Robert Schuman Centre for Advanced Studies at the European University Institute, working in the framework of the 'BlockchainGov' ERC Project. His research focuses on blockchain governance and, more broadly, on cooperative governance. He received his Ph.D. from Leiden Law School, Leiden University for his dissertation entitled: *The Emergence of Democratic Firms in the Platform Economy: Drivers, Obstacles and the Path Ahead*. He has published several articles in academic journals such as *Policy & Society*, *Ondernemingsrecht*, *Georgetown Law Technology Review*, *Technology and Society*, *Topoi* and *Erasmus Law Review* on topics pertaining to blockchain governance and the formation of a nascent type of cooperative business: platform cooperatives.

He is currently co-authoring a book on blockchain governance with Dr. Primavera de Filippi and Dr. Wessel Reijers, under contract with a leading academic press. He is also editing a blockchain reader, containing primary materials and commentary from key developments and moments in the history of blockchain, with Dr. Primavera de Filippi and several colleagues. As a corporate law researcher, he has earlier published a book *Freedom of Establishment for Companies in Europe (EU/EEA)* with his PhD supervisor, Iris Wuisman. Morshed is a Research Affiliate of the Institute for the Cooperative Digital Economy at The New School in New York City, a CLARITY International Fellow of NCBA Clusa International, and has been called to the Bar of England & Wales and Bangladesh. He has also acted as a consultant on matters regarding decentralized autonomous organizations, and cooperative law and governance for the International Cooperative Alliance and NCBA Clusa International, and as an expert for the UN Department of Economic and Social Affairs, the OECD, the European Commission, as well as several local and national government bodies

Dr. Patricia Boshe is a data protection trainer, researcher and consultant. She is a co-founder and co-director of the African Law and Technology Institute (AFRILTI); a research institute focusing on the interrelation between law, technology and society from an interdisciplinary perspective. She has more than 10 years' experience as a law lecturer in Tanzania. Currently, a senior researcher at the Research Centre for Law and Digitalization (FREDI) at the University of Passau in Germany. Some of her research activities involve assessments and critiques on privacy and data protection in Africa. Her publication record includes a book on data protection, book chapters and over dozen of international referred journal articles, book reviews and practical legal comments.

Dr. Courtney Radsch is scholar practitioner working at the nexus of technology, media, and rights. Her work focuses on tech policy, platform governance, AI and influence operations, and media sustainability. Radsch writes and speaks frequently about these issues and has testified before Congress, participated in expert consultations at the United Nations, EU, OSCE, and OECD, and provided expertise to technology platforms on policy and product design and impact. She has led advocacy missions and media assessments in more than a dozen countries and trained journalists and activists around the world. She is a post-doctoral fellow at UCLA's Institute for Technology, Law and Policy and a fellow at the Center for Democracy and Technology, the Center for International Governance Innovation, and the Media and Journalism Research Center. She has worked as a journalist in the Middle East, an international diplomat with UNESCO, and a human rights advocate with leading organizations including the Committee to Protect Journalists, Freedom House and ARTICLE 19. She is a founding member of the ACOS (A Culture of Safety) Alliance for journalist safety and the Christchurch Call Advisory Network, and works on responsible tech and platform accountability, content moderation, and countering violent extremism online as a member of the Global Network Initiative (GNI), the International Science Council's Panel of Experts, and the Global Internet Forum to Counter Terrorism (GIFCT) International Advisory Committee. Her book *Cyberactivism and Citizen Journalism in Egypt: Digital Dissidence and Political Change* was published in 2016 and she holds a Ph.D. in international relations.

Mike Ananny is an Associate Professor of Communication and Journalism at the University of Southern California's Annenberg School for Communication and Journalism, where he studies the intersection of journalism practice and technology design, the public significance of digital news infrastructures, and the ethics of algorithmic systems. He is also on the Steering Committee of USC's Science, Technology and Society research cluster, a Faculty Fellow with USC's Society of Fellows in the Humanities, and co-directs the interdisciplinary research group "MASTS" (Media as SocioTechnical Systems) and the Sloan Foundation project Knowing Machines (with Kate Crawford and Jason Schultz). He was a 2022 Visiting Professor at the University of Helsinki Institute for Social Sciences and Humanities, a 2018-19 Berggruen Fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford University, and has held fellowships and scholarships with the Columbia University's Tow Center for Digital Journalism, Harvard's Berkman-Klein Center on Internet and Society, Stanford's Center on Philanthropy and Civil Society, the Pierre Elliott Trudeau Foundation, LEGO, and Interval Research. He was a founding member of Media Lab Europe, a postdoc with Microsoft Research's Social Media Collective, and has consulted for LEGO, Mattel, and Nortel Networks. His PhD is from Stanford University (Communication), SM from the MIT Media Lab (Media Arts & Sciences), and BSc from the University of Toronto (Human Biology & Computer Science). He has published in various academic and popular venues, is the author of *Networked Press Freedom* (MIT Press, 2018), co-editor (with Laura Forlano and Molly Wright Steenson) of *Bauhaus Futures* (MIT Press, 2019), and is preparing a manuscript on the public power of silence and mediated absences (under contract with Yale University Press). He has written for popular press publications including *The Atlantic*, *Wired Magazine*, Harvard's *Nieman Lab*, and the *Columbia Journalism Review*.

Saule Omarova is the Beth and Marc Goldberg Professor of Law at Cornell University and a Senior Fellow at Roosevelt Institute. Her scholarship focuses on systemic risk regulation, financial technology, and structural trends in global financial markets. Prior to joining academia, she practiced banking law at a premier New York law firm, and served at the U.S. Treasury Department as Special Advisor for Regulatory Policy to the Under Secretary for Domestic Finance. In 2021, Professor Omarova was President Biden's nominee for the U.S. Comptroller of the Currency. She holds a Ph.D. degree from the University of Wisconsin-Madison and a J.D. from Northwestern University.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Economist Intelligence Unit (2023). Democracy Index 2022. Frontline Democracy and the Battle for Ukraine. London: EIU.

2 DataReportal (2022), "Digital 2022 Global Digital Overview," retrieved from https://datareportal.com/reports/digital-2023-global-overview-report .

3 Grand View Research (2023). "Artificial Intelligence Market Size, Share & Trends Analysis Report By Solution, By Technology (Deep Learning, Machine Learning), By End-use, By Region, And Segment Forecasts, 2023 – 2030", retrieved from https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market .

4 U.S. Department of State Press Release: "Secretary Antony J. Blinken With Maria Ressa on Digital Diplomacy and Human Rights Online", (Washington D. C, June 7 2022).

5 White House, National Cybersecurity Strategy, (Washington D. C., March 2023), p. 3.

6 Rebecca MacKinnon, "Liberation Technology: China's "Networked Authoritarianism"," Journal of Democracy 22, no. 2 (2011): 32-46, p. 33.

7 Alina Polyakova and Chris Meserole, "Exporting Digital Authoritarianism: The Russian and Chinese Models", (Washington D. C: Brookings Institute, 2019).

8 Adrian Shahbaz, "Freedom on the Net 2018: The Rise of Digital Authoritarianism, (Washington D. C: Freedom House, 2018), p. 1.

9 Erol Yayboke and Sam Brannen, "Promote and Build: A Strategic Approach to Digital Authoritarianism". (Washington D. C: CSIS, 2020), p. 1

10 Monique Taylor, "China's Digital Authoritarianism: A Governance Perspective", (Cham: Palgrave Macmillan, 2022); Elina Sinkkonen and Jussi Lassila, "Digital Authoritarianism in China and Russia: Common Goals and Diverging Strandpoints in the Era of Great-power Rivalry, (Helsinki: Finnish Institute of International Affairs, 2020).

11 Marc Owen Jones, "Digital Authoritarianism in the Middle East: Deception, Disinformation and Social Media", (Oxford: Oxford University Press, 2022).

Part I
Digital Geopolitics

In this section:

- **Whose Internet? Authoritarianism and the Struggle Over Governance**
- **Foreign Actors and Digital Authoritarianism in Africa: Recent Trends on Methods and Their Human Rights Impacts**

Berggruen Institute | Canada | UCLA Institute for Technology, Law & Policy | USC Dornsife Center on Science, Technology, and Public Life

Whose Internet? Authoritarianism and the Struggle Over Governance

Kal Raustiala
UCLA*

I. The Internet's Origins

The Internet's origins in California—the very first message was sent from UCLA to Stanford in October, 1969-- built a set of liberal and even libertarian elements into Internet governance from its debut. The Internet remained largely Californian, if increasingly American, for many years. With the birth of the first web browsers in the 1990s the Internet escaped its university and government settings and became a widely-adopted public good and commercial and social space. Over the last two decades the Internet fully globalized; by 2018 more than half the world was online.[1]

As the Internet has dramatically transformed, authoritarian regimes—fearful of its implications for social movements and political organizing--have in turn increasingly sought to transform it. The Internet was widely extolled as a force for openness in the 2000s and early 2010s; indeed, as a "liberation technology."[2] The reality has proved far more mixed. Authoritarian regimes, most notably China, have sought to clamp down on the original vision of a free and open Internet, and have done so with surprising success. They have increasingly cabined and controlled the Internet at home, often turning it into a powerful tool of surveillance and repression.[3]

Externally, a sharper and less successful struggle has ensued. In forums ranging from the United Nations to the International Telecommunications Union (ITU) to the Internet Corporation for Assigned Names and Numbers (ICANN), the battle has been over who rules the Internet and how they rule. This competition between, broadly, a Californian and Chinese vision of the Internet—between liberal and authoritarian governance of the digital—has dominated 21st century debate. Yet it has not fully played out as most theories of authoritarian international law expect.

Sovereignty and stability have, as theories of authoritarianism expect, been central goals of Internet authoritarians.[4] These regimes have increasingly extolled state control of digital flows. They have also stressed the importance—in an important and perhaps dangerous area of tangency with the West—of digitally-derived and delivered "harms." In extremis, they have even shut down the Internet.[5]

Yet rather than seek thinner forms of cooperation and looser global arrangements, authoritarian regimes have sought stronger global rules via delegation to multilateral organizations in which they have an important voice. Conversely, liberal democracies, in particular the U.S., have resisted these efforts. They have sought to limit and even avoid multilateralism. They instead attempted to entrench multistakeholderism: governance in which non-state actors rule and regulate alongside state representatives.[6] These unusual

dynamics reflect not only the particular trajectory of the Internet but also the central role played by private actors in the digital domain from its very origins in mid 20th century California.

II. International Law and Internet Governance

International law is facially neutral as to regime type. But in practice, at least since the Second World War international law has largely been dominated by democracies and often pushed forward, in its evolution, by the advanced industrial democracies of the OECD world. In the 21st century that tendency has abated and a counter-trend, of authoritarian international law, has emerged with greater force. To be sure there were always powerful authoritarian states with a major role in international law, not least the Soviet Union, a state with, among other roles, a permanent seat on the UN Security Council. But substantial aspects of international law, in particular economic law and to a degree as well human rights law, remained largely outside the Soviet reach and were predominantly forged by the West.

Authoritarian international law has been prominently defined by Tom Ginsburg "as legal rhetoric, practices, and rules specifically designed to extend the survival and reach of authoritarian rule across space and/or time."7 Much of the rise of authoritarian international law can be ascribed to the rise of China over the past 25 years as a central player in world politics. But it is not limited to China. In the digital domain (and elsewhere) China has been joined by like-minded states, such as Russia, Iran, and Turkey, who seek to impose greater control over what they perceive as a dangerously open and free Internet. Domestically, China's "Great Firewall" and a host of related efforts have shown President Bill Clinton's infamous claim that controlling the Internet was akin to nailing Jello to the wall to be wholly false.[8] Indeed, in recent years the Internet has become, in many respects, a tool of authoritarian governance, enabling regimes to maintain power, control, and legitimacy at great scale.[9]

At the global level this has led to a deep conceptual and political divide over Internet governance. The early Internet was governed and generally seen as almost outside the state or traditional governance. It was populated and run by hobbyists, academics, and computer scientists. John Perry Barlow's (in)famous 1996 "declaration of independence," a paean against sovereignty in cyberspace, was utopian but emblematic of the (rapidly waning) era of the Internet's inception.[10] The early Internet was also strongly American. Virtually every early advance, every node, every theory of the Internet in the early days reflected its American academic roots--and federal DARPA funding. Even as the Internet became a commercial and popular platform in the 1990s the basic structure remained largely in private hands. This strongly private orientation continues as a practical matter today: the hardware of the Internet is largely not in state hands, nor are the key software and applications. The goal of the United States government was, and remains, an open and free Internet; an analogue to the vision of interdependent globalization the Clinton administration advocated in so many policy domains.

As governance of the Internet moved from a quasi-volunteer system to more regularized regulatory institutions, this tendency remained. ICANN, for example, was deliberately funded by the US Commerce Department as a way to spin off regulatory activities from the federal government and—most importantly—insulate them from the now-looming specter of greater multilateral control. Formed in 1998 and headquartered in California, ICANN

was in part a reaction to early attempts by some governments to use existing international institutions, in particular the International Telecommunications Union, to govern what was now becoming a crucial global resource. The ITU had a strong claim to play a greater role with regard to digital communication. But for the US in particular, the ITU, with its one nation, one vote system, was too easily dominated by the authoritarian, the authoritarian-leaning, and the statist governments of the world.

Efforts at greater state control continued in the early 21st century. In the fall of 2011, for example, India issued a call to "place Internet governance under the auspices of the UN, or, as some have characterized it, 'in a box with a UN label stamped on the side.'"[11] Shortly after, the OECD countered with a Communique on Principles of Internet Policy-Making that expressly endorsed multistakeholderism over multilateralism. "Due to the rapidly changing technological, economic and social environment within which new policy challenges emerge," the OECD declared, "multi-stakeholder processes have been shown to provide the flexibility and global scalability required to address Internet policy challenges."[12]

By the 2012 World Conference on International Telecommunications the struggle over which vision of the Internet-- the Californian or Chinese—and, relatedly, which model of governance—the multilateral or the multistakeholder-- would prevail came to a head. Held in the United Arab Emirates, the assembled states quarreled over how much control over "international telecommunications" (read, the Internet) to grant governments. The US delegation, refusing to sign a final text that it saw as yielding too much control to states, authoritarian or otherwise, declared that it would not support a treaty "that is not supportive of the multistakeholder model."[13]

The US was not alone in this view; many liberal democracies opposed, and feared, the move toward greater state control that was being increasingly pushed in global forums. French president Emmanuel Macron, in 2018, explained this view in exactly the terms laid out here. "To be very politically incorrect," Macron declared,

> We are seeing two types of Internet emerge...there is a Californian form of Internet, and a Chinese Internet. The first is the dominant possibility, that of an Internet driven by strong, dominant, global private players, that have been impressive stakeholders in this development, that have great qualities and with which we work, but which at the end of the day are not democratically elected. . . . On the other side, there is a system where governments have a strong role, but this is the Chinese-style Internet: an Internet where the government drives innovations and control.[14]

The UN itself increasingly has become the site of authoritarian governance initiatives. In late 2019, for instance, a resolution sponsored by a number of authoritarian states—Belarus, Burma, China, Russia, Venezuela, among others—established a new working group on cybercrime cooperation. The goal is a new treaty. But for many, the proposed treaty proposed is not really aimed at cybercrime, but rather at "who controls the Internet."[15]

III. Multilateralism and Multistakeholderism

There is substantial evidence that authoritarian states behave differently than non-authoritarian states. In assessing the influence of regime type on international law, Ginsburg argues that:

> We should expect, *ceteris paribus*, less willingness to include broad third-party dispute resolution clauses in treaties, and shallower legal commitments with more flexibility. We should also expect authoritarians to be less interested in public visibility, both in the sense of making fewer public binding commitments, and being less willing to tolerate institutions that increase domestic transparency.[16]

Is this consistent with the empirical record in Internet governance? The answer is mixed. Authoritarian states have been more willing, not less, to rely on multilateral institutions to create binding commitments over Internet governance. They have resisted flexibility. They have sought stronger rules. They have generally wanted to see international law deployed as a tool of regulation and embedded in established, state-driven institutions with substantial bureaucratic capacity (such as the ITU). In these respects, the Internet seems to offer a marked contrast to what is normally predicted for authoritarian international law.

On the other hand, the Internet governance debate illustrates other points commonly associated with authoritarian international law. Authoritarian states do seek to use international institutions to promote autocracy and they do aim to protect the state from the destabilizing influences borne by digital communications. They also seek to promote "collective" rights over individual freedoms.

Yet on the largest points—should multilateralism be the chief form of governance; should international organizations play a key regulatory role; should unanimity or majority rule prevail—the position of the major authoritarian states fits at best uneasily, and often largely at odds, with authoritarian international law's predictions. And it is the liberal democracies, led by the U.S., who oppose multilateral governance vociferously.

This rejection by the West of multilateralism and the embrace of multistakeholderism can be explained in several ways. Multistakeholderism has normative appeal for many states precisely because it allows a wide range of actors, public and private, into the circle of influence and decision making. For societies with a strong commitment to a vibrant public sector, this can seem a natural progression. As a result here is increasing attention to multistakeholderism today in a number of areas of global governance, such as global health (see e.g. the Global Fund; GAVI).

This move underscores a growing concern on the part of some governments with traditional multilateral approaches to international law. And it reflects a perhaps shrewd political calculation. The policy preferences of powerful governments—in particular, the advanced industrial democracies that are also the home of many well-resourced firms and NGOs—might be best realized indirectly through greater incorporation of a wide variety of private sector actors, rather than directly through traditional state-centric international law models.

And as political power increasingly disperses in the world, the appeal of traditional multilateralism is likely to diminish still further for the U.S. and its allies—and not simply due to the idiosyncratic and unreliable approach of the former (and perhaps future) Trump administration. In some settings, such as the UN Security Council, entrenched rules continue to favor traditional great powers. But many international legal bodies operate on a one nation-one vote system and are subject to increasing demands for inclusive leadership. The greater inclusion of private actors in international organizations that

has marked the last several decades is less a sign of governments ceding power than a sign that NGOs, firms, and the like bring valuable resources to the regulatory table.[17] But in the vast majority of settings, participation by non-state actors is limited to voice but not decisions.

What distinguishes multistakeholderism, and perhaps makes it increasingly appealing to powerful Western governments who foresee greater power dispersion, is precisely that it goes much further. By granting non-state actors direct governance roles, and ensuring that states cannot dominate global rulemaking, multistakeholder governance limits the ability of authoritarian states to steer governance in their direction. And with regard to the Internet, it does even more. The libertarian, California Internet of the past is indeed in the past. But many of the key Internet governance actors remain committed to such an ethos, and of course many of the most critical global firms that manage and dominate the Internet are Californian. That constellation of interests neatly allows the liberal democracies to achieve many of their preferences without directly wielding their power.

None of this is to suggest that authoritarian states cannot control the Internet. Indeed, they do already. But at the global level that control is inherently limited. The entrenching of multistakeholder approaches over multilateral approaches—while perhaps not necessarily sustainable over the long term—helps to extend those limitations on authoritarianism, even in the face of rising coordination and power on the part of the authoritarian world.

IV. Emergent Compromise?

The balance between multilateralism and multistakeholderism remains hotly contested. But the history of governance in the digital domain suggests multistakeholderism can operate effectively, even as security and other concerns complicate its future. Of course, classical economists thought much the same about the international economy and the trading system in the pre First World War era. If only certain limited commitments were made, such as no tariffs, use of the gold standard, and the like—the system would be allegedly self-regulating. But politics intruded to disrupt this system, which never really operated without state action anyway. And by the mid-20th century, after the world was repeatedly wracked by war and depression, the compromise of embedded liberalism, in the late John Ruggie's famous phrase, took hold.[18]

Perhaps the same broad process is occurring today in the digital domain, as witnessed by increasing calls for regulation of global technology firms across the West. If so, over time some kind of embedded digital openness may take root, in which the Internet remains a single, broadly cohesive network, but with many national level controls to rein in the worst excesses of unfettered digital flows. Arguably the move to greater content regulation and moderation, the many restrictions on data and privacy being promulgated around the world, and new efforts to devise shared norms for state operations in cyberspace are evidence of such an emergent compromise, one pioneered by the European Union and its unparalleled regulatory powers.[19]

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 https://ourworldindata.org/internet

2 Larry Diamond and Marc Plattner, eds, Liberation Technology (JHU Press, 2012).

3 S. Kalathil, Beyond the Great Firewall: How China Became a Global Information Power. Washington, DC: Center for International Media Assistance, 2017.

4 Tom Ginsburg, Authoritarian International Law? American Journal of International Law, 114, 2 (April 2020).

5 https://www.cnn.com/2019/12/21/asia/internet-shutdowns-china-india-censorship-intl-hnk/index.html

6 See, e.g., Mark Raymond and Laura DeNardis, "Multistakeholderism: Anatomy of an Inchoate Global Institution," Int'l Theory 7, (2015) pg 572.

7 Ginsburg, supra note 4, at 228.

8 https://www.nytimes.com/2000/03/09/world/clinton-s-words-on-china-trade-is-the-smart-thing.html

9 Gerschewski, J., & Dukalskis, A. (2018). How the Internet Can Reinforce Authoritarian Regimes: The Case of North Korea. Georgetown Journal of International Affairs, 19, 12–19.

10 https://www.eff.org/cyberspace-independence.

11 Joe Waz & Phil Weiser, Internet Governance: The Role of Multistakeholder Organizations, 10 Journal of Telecommunications Technology Law 331 (2012).

12 Communique on Principles for Internet Policy-Making: The Internet Economy: Generating Innovation and Growth, Org. for Econ. Coop. & Dev. 4 (2011), https://www.oecd.org/digital/ieconomy/48387430.pdf.

13 Eric Pfanner, U.S. Rejects Telecommunications Treaty, N.Y. Times, Dec. 14, 2012, at 1.

14 Emmanuel Macron, President, France, 2018 Speech by French President Emmanuel Macron to the Internet Governance Forum (2018), https://www.intgovforum.org/multilingual/content/igf-2018-speech-by-french-president-emmanuel-macron (emphasis added).

15 Tom Ginsburg, Democracies and International Law (Cambridge 2021) at 228.

16 Ginsburg, supra note 4 at 231.

17 Kal Raustiala, States, NGOs, and International Environmental Institutions, International Studies Quarterly, 41, 4 (Dec. 1997).

18 John Gerard Ruggie, International Regimes, Transactions, and Change: Embedded Liberalism in the Postwar Economic Order, International Organization 36, 2 (Spring 1982).

19 Anu Bradford, The Brussels Effect: How the EU Rules the World (Oxford, 2020).

Foreign Actors and Digital Authoritarianism in Africa: Recent Trends on Methods and Their Human Rights Impacts

Tomiwa Ilori
University of Pretoria*

I. Foreign Actors and Digital Authoritarianism in Africa

Digital authoritarianism now involves foreign actors in Africa. These actors work with African state actors to facilitate various digital authoritarian practices such as networked surveillance and information disorder. For example, the Chinese government supply African governments who have questionable human rights records with surveillance technologies while Russian entities plant troll farms to influence elections.[1] Israeli-based companies also sell privacy-intrusive spyware to African governments while a UK-based surveillance technology company once provided surveillance infrastructure to Uganda without ensuring adequate human rights safeguards are in place.[2] While these examples primarily show that authoritarianism has been reinvented in African countries for the digital age,[3] they also show that it transcends borders and requires foreign actors to thrive. However, there are limited conversations on how to build resistance against these new methods of digital authoritarianism. Therefore, this contribution identifies how various stakeholders can resist these methods and trends of digital authoritarianism enabled by foreign actors.

II. Methods and Recent Trends in Africa

Digital authoritarianism has been described as the pervasive use of technologies by state actors 'to surveil, repress, and manipulate domestic and foreign populations.'[4] It has also been identified as 'a new global force of disruption.'[5] Deibert classifies digital authoritarian methods into four categories.[6] The first-generation category are methods of information controls whereby domestic actors raise national cyber-walls to restrict their citizens' access to global information infrastructure e.g. Great Firewall of China. This method has also been adopted in countries like Iran, Pakistan, Saudi Arabia and Vietnam. The second-generation category is characterized by using laws and regulations to compel private sector actors to provide backdoor software, facilitate state-ordered censorship and comply with surveillance requests. African countries like Zimbabwe, Ethiopia, Nigeria, Tanzania, Uganda, Zambia, Mozambique and Cameroon all have laws requiring private sector actors to provide for backdoor tools to access citizens' data.[7] The third-generation method involves targeted espionage and surveillance thereby becoming more offensive rather than being defensive like in the first- and second-generation categories. Recently, Rwanda

was reported to have put the phone of South African President, Cyril Ramaphosa on a list of people to be targeted by Israeli-made Pegasus spyware.[8] The last category, which emerged in the past decade, is the assertion of one or two of these methods at the international level by foreign actors that seek to export their authoritarian methods worldwide. This latest method has also found its foothold in African countries.

Considering that digital authoritarianism also involves foreign populations and given Deibert's fourth categorization, it is without doubt that its practice requires foreign actors in African contexts. While the influence and interests of foreign actors in African countries take many forms, as far as it relates to digital authoritarianism in African countries, these influence and interests are noticeable in two major trends: networked surveillance and information disorder.[9]

Networked Surveillance

According to the Business Research Company, the global surveillance infrastructure market is currently estimated at US$130.08 billion.[10] Most surveillance companies are concentrated in Europe with 23 out of 27 European Union countries having surveillance companies.[11] The United Kingdom has the most companies followed by France, Germany, Italy and Sweden. These companies provide various surveillance services including setting up surveillance infrastructure and selling surveillance tools to other countries. Israel, the United States, Russia and China are also renowned for their surveillance companies.[12] These technologies, which include GPS tracking devices, video area persistent surveillance and commercial spyware have been purchased from some of these countries and have been indiscriminately deployed by African governments under the guise of ensuring national security.[13]

In particular, there have been debates about the rise of Chinese surveillance technologies and what it means for human rights and democracy in Africa.[14] Currently, China's investments in digital infrastructure in Africa is mainly driven by its Digital Silk Road (DSR) initiative.[15] The DSR initiative is China's socio-economic narrative that focuses 'solely on connectivity and the digital economy.'[16]

Fig 1:

Source: Deutsche Welle, 2019[17]

Some of the projects of the DSR include Safe City (Huawei) and Smart City (ZTE) projects involve the deployment and use of Chinese surveillance technologies.[18] These projects are sold to African governments based on the claim that they will help curb crimes but this is not usually the case as crime rates have actually increased in most places Huawei technologies have been deployed and used in African countries.[19] African countries in partnership with Huawei to develop their Smart City project include Algeria, Botswana, Côte d'Ivoire, Egypt, Ethiopia, Ghana, Kenya, Mauritius, Morocco, Nigeria, Rwanda, South Africa, Uganda, Zambia and Zimbabwe while

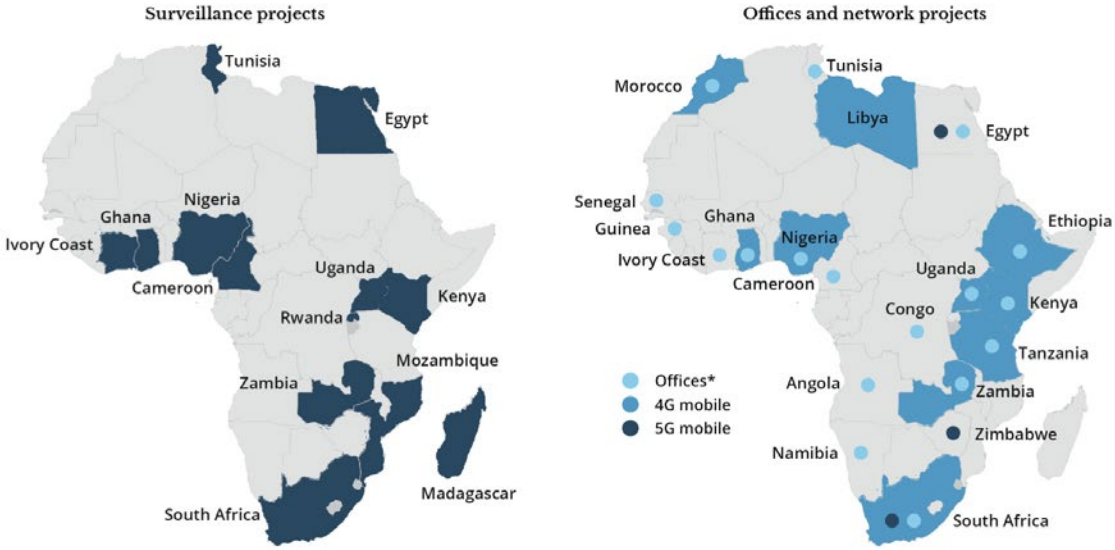
ZTE has a smaller presence, with projects launched in only five countries: Egypt, Ethiopia, Nigeria, Sudan and Zambia.[20] In addition to this, while the DSR initiative as a whole may have positive overall impacts, they are still used by African governments to suppress dissent in countries like Ethiopia, Uganda and Zambia.[21] The DSR initiative and surveillance technologies in Africa also obscure another important dynamic in China-Africa relations – the Sino-African economic relationship.[22]

Between 2003 and 2020, foreign direct investment from China to African countries was estimated at US$49 billion while as at 2018, China's total lending to African governments and their State-Owned Enterprises (SoEs) is estimated at US$148 billion.[23] Given this relationship, most African governments may be willing to look the other way when it comes to assessing China's human rights records or ensuring that human rights are complied with considering African governments also have questionable human rights records themselves.[24]

In addition to China, Israeli-based surveillance companies have also been linked to African countries. For example, in a study by Citizen Lab, Pegasus, a spyware has been found on internet service providers in Togo, Algeria, Morocco, Cote D'Ivoire, Tunisia, Kenya, Rwanda, South Africa, Uganda and Zambia.[26] The Ethiopian government has also engaged the services of surveillance companies in Italy, the United Kingdom and Germany.[27]

It is important to note that the focus on Chinese surveillance technologies should also not obscure the fact that other countries in Europe discussed above are also part of the global surveillance industry.[28] It is also important to note that while China's human rights records may be worrisome, it remains unclear how or whether it is foisting its surveillance methods on African governments.[29] Such influence might be difficult to justify because it absolves African governments of any responsibility to resist such technologies.

Fig 2: China's surveillance-related projects in Africa



Source: The Wall Street Journal, 2019[25]

The methods and recent trends above point to one major issue: European countries, China, Israel and others with surveillance companies deal with African governments when it comes to enabling surveillance practices. Therefore, notwithstanding the domestic democratic values that exist in Western or Eastern countries, they still facilitate these practices which are also rights-averse in African countries. This relationship suggests that foreign actors, whether state or non-state, are a network of surveillance actors that facilitate digital authoritarianism in African countries.

Information Disorder

Foreign influence efforts promoting various forms of information disorder have targeted at least 24 countries across the world since 2013.[30] These disorder manifests mostly through disinformation campaigns and at least 72% of these efforts were made by Russia, China, Iran and Saudi Arabia.[31]

In African countries, these efforts are often deployed during elections and they are mostly geared towards pro-Russian conversations and alignment of domestic actors towards Russian interests.[32] These efforts have been ongoing in at least 20 African countries where disinformation campaigns were deployed to distort information on various local and international issues. Oftentimes, these efforts are carried out by proxy through local organizations in African countries. For example, in 2020, social media platforms uncovered Russia-led troll farms in Nigeria and Ghana targeting the United States,[33] while more recently, it was noted that Russian disinformation campaigns helped drive French forces out of Burkina Faso.[34] China has also been identified to spread some of these efforts across African countries.

In addition to this, a report by Stanford Internet Observatory analyzed 73 Facebook pages linked to Russian foreign efforts in six countries namely Libya, Sudan, the Central African Republic, Madagascar, Mozambique, and the Democratic Republic of the Congo. The report noted that these efforts were focused around Russian-aligned actors and politicians through impersonation and propaganda.[35]

III. Human Rights Impact

Digital authoritarianism has raised many concerns including its impacts on human rights in Africa.[36] Digital authoritarian methods such as networked surveillance and information disorder as facilitated by foreign actors for African governments pose huge risks to digital freedoms such as online expression, privacy, association and assembly and various other rights.[37] For example, in 2011, surveillance technologies supplied by a UK-based company to the Ugandan government were used to target political opposition and the media.[38] This has also been the case in Ethiopia, Nigeria, Zimbabwe, and many other countries where surveillance technologies are purchased and deployed for extra-legal practices without transparency or accountability.[39]

Oftentimes, governments' basis for these practices is to ensure national security or public order.[40] However, in most instances, these bases are neither to ensure national security nor public order but to satisfy and sustain the whims of autocratic governments.[41] These instances are particularly worrisome especially given how African governments have now mandated digital identification for their citizens. Currently, at least 48 African governments have a national identity (national ID) system which involves collection of personal information in exchange for accessing public services. Some of these services include accessing public amenities, immigration services, voting during elections etc.

Due to the collection, processing and storage of personal information of individuals that is usually involved to roll out national IDs, it has become necessary that such systems are used only when adequate human rights safeguards are in place.[42] This is because African governments collect a lot of personal information about their citizens but there are usually no adequate human rights safeguards for such data. For example, out of all the 48 African countries with national ID systems, only 20 countries have data protection laws and functional data protection agencies (DPAs) to implement such laws.[43]

Information disorder is also known to have negative human rights impacts. These impacts may be seen in their potential to reduce public trust in the internet as a public good and how it has surreptitiously manipulated citizens' rights to hold opinions that could inform their social and political views and choices.[44]

Taking all these methods by foreign actors together, the digital civic space also faces existential threats as digital authoritarianism seeks to exert state control over its citizens. Online publics are unable to network without circumvention technologies and are constantly faced with threats to their lives. It is also important to note that most African governments assisted by these foreign actors have been known to harass and threaten vulnerable groups like migrants, sexual minorities, and political prisoners in Africa.[45]

These analyses show an obvious pattern that foreign actors facilitate digital authoritarianism in African democracies and this violates human rights, threatens the digital civil space and makes the vulnerable even more vulnerable. However, the responsibility on how to resist these impacts need to accommodate more than just foreign and domestic actors but also include broader international actors.

IV. Resistance Strategies

From the methods, trends and impacts of foreign actors in digital authoritarian practices in Africa, one point is clear, there is need for collaboration and capacitating at-risk groups. This is because digital authoritarian practices such as networked surveillance and information disorder primarily maximize technologies for pervasive control, therefore counter-measures that seek to resist must be dynamic, be ready to wrestle such control from foreign actors and ensure that African governments obey the rule of law in their digital policies and practice.

One means of countering digital authoritarian practices in African countries is through collaborative efforts among governments, civil society (domestic and international), academia, at-risk groups, international and regional institutions. While this is a difficult task, it is possible. This is because given the extra-territorial nature of digital authoritarianism today, neither domestic nor international counter-measures alone can work. There is need for stakeholders to network and commit to countering these practices as foreign and domestic actors themselves. One goal for collaboration may be to research these methods, trends and impacts even far beyond this contribution to help inform their strategies. For example, there is need for standard-setting when it comes to the role foreign actors play in facilitating digital authoritarianism practices in African countries. Foreign actors also need to be made more responsible by ensuring rigorous due diligence practices including human rights impact assessments of surveillance technologies before they are deployed in any African country. This will involve making human rights protection as a condition precedent for engaging in such relationship with African governments and not an after-thought.

Another means of countering such measures is through collaborative advocacy on transparent and accountable foreign actors. Purchase, transfer and use of surveillance technologies must be subject to radical transparency practices that provide basis for such technologies and makes them accountable where harm arises as a result of their use. In this instance, advocacy should be directed at both foreign and domestic actors in order to spotlight how networked authoritarianism impacts human rights and importantly, why they must be held accountable.

In addition, in many instances where foreign actors facilitate digital authoritarian practices in African countries, the rights of civil society actors and vulnerable groups have been most impacted. Given the tough work of resisting such practices, it is equally important that these actors and groups are kept safe. Such safety can be ensured by providing various digital security measures for at-risk civil society actors and vulnerable groups. These measures include digital security and safety training, providing digital security tools and practicing consistent digital hygiene.

V. Decoding Foreign Actors in Digital Authoritarian Practices

In order to properly decode digital authoritarianism in African countries, it is important to identify the various actors involved in its practice. This contribution identifies foreign actors including international governments and businesses as facilitators of authoritarianism in African countries. It identifies the various methods involved, the current trends and the impacts of these trends on human rights in Africa. It notes that in the course of facilitating digital authoritarianism for African governments, foreign actors indirectly violate human rights and this calls for counter-measures from various stakeholders. It concludes that such counter-measures must be collaborative and capacitate at-risk actors and groups to resist networked surveillance and information disorder as facilitated not just by foreign actors but also by domestic actors in African countries.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Robert Morgus, "The Spread of Russia's Digital Authoritarianism," Artificial Intelligence, China, Russia, and the Global Order (Air University Press, 2019), https://www.jstor.org/stable/resrep19585.17.

2 Bill Marczak et al., "Hide and Seek: Tracking NSO Group's Pegasus Spyware to Operations in 45 Countries" (University of Toronto, September 2018), https://tspace.library.utoronto.ca/bitstream/1807/95391/1/Report%23113--hide%20and%20seek.pdf.

3 Marie Lamensch, "Authoritarianism Has Been Reinvented for the Digital Age," Centre for International Governance Innovation, July 9, 2021, https://www.cigionline.org/articles/authoritarianism-has-been-reinvented-for-the-digital-age/.

4 Alina Polyakova and Chris Meserole, "Exporting Digital Authoritarianism: The Russian and Chinese Models" (Brookings, August 2019), https://www.brookings.edu/wp-content/uploads/2019/08/FP_20190827_digital_authoritarianism_polyakova_meserole.pdf.

5 Lamensch, "Authoritarianism Has Been Reinvented for the Digital Age."

6 Ron Deibert, "Cyberspace Under Siege," Journal of Democracy 26, no. 3 (2015): 65–70, https://doi.org/10.1353/jod.2015.0051.

7 CIPESA, "Digital Authoritarianism and Democratic Participation in Africa," June 2022, https://cipesa.org/wp-content/files/briefs/Digital-Authoritarianism-and-Democratic-Participation-inå-Africa-Brief-.pdf.

8 Carien du Plessis, "Pegasus spying scandal: Rwanda targeted South Africa's Ramaphosa," July 2021, https://www.theafricareport.com/111202/pegasus-spying-scandal-rwanda-targeted-south-africas-president-ramaphosa/

9 Marlies Glasius and Marcus Michaelsen, "Illiberal and Authoritarian Practices in the Digital Sphere," International Journal of Communication 12 (2018): 3800–3804, https://ijoc.org/index.php/ijoc/article/view/8899.\\uc0\\u8221{} {\\i{}International Journal of Communication} 12 (2018

10 The Business Research Company, "Surveillance Technology Global Market Report," January 2023, https://www.thebusinessresearchcompany.com/report/surveillance-technology-global-market-report.

11 Privacy International, "The Global Surveillance Industry," July 2016, https://privacyinternational.org/sites/default/files/2017-12/global_surveillance_0.pdf.

12 Privacy International.

13 CIPESA, "Digital Authoritarianism and Democratic Participation in Africa."

14 Willem Gravett, "Digital Neo-Colonialism: The Chinese Model of Internet Sovereignty in Africa," African Human Rights Law Journal 20, no. 1 (2020): 125–46, https://doi.org/10.17159/1996-2096/2020/v20n1a5.

15 David Gordon, "The Digital Silk Road: Introduction," IISS, December 6, 2022, https://www.iiss.org/blogs/analysis/2022/12/digital-silk-road-introduction; Motolani Agbebi, "China's Digital Silk Road and Africa's Technological Future" (Council on Foreign Relations, February 1, 2022), https://www.cfr.org/sites/default/files/pdf/Chinas%20Digital%20Silk%20Road%20and%20Africas%20Technological%20Future_FINAL.pdf."

16 Gordon, "The Digital Silk Road."

17 Chiponda Chimbelu, "China Leads Investments in African Tech Infrastructure," Deutsche Welle, May 3, 2019, https://www.dw.com/en/investing-in-africas-tech-infrastructure-has-china-won-already/a-48540426.

18 Iginio Gagliardone, "The Impact of Chinese Tech Provision on Civil Liberties in Africa" (South African Institute of International Affairs, December 2020), 13, https://saiia.org.za/wp-content/uploads/2020/12/Policy-Insights-99-gagliardone.pdf.

19 Gagliardone, 16.

20 Gagliardone, 13.

21 Agbebi, "China's Digital Silk Road and Africa's Technological Future," 11.

22 Justin Bryant, "Africa in the Information Age: Challenges, Opportunities, and Strategies for Data Protection and Digital Rights," Stanford Technology Law Review 24, no. 2 (2021): 426–30, https://law.stanford.edu/wp-content/uploads/2021/05/BryantAfricaInTheInformationAge.pdf.

23 Deborah Brautigam, Yufan Huang, and Kevin Acker, "Risky Business: New Data on Chinese Loans and Africa's Debt Problem" (Washington, DC: China Africa Research Initiative (CARI), School of Advanced International Studies (SAIS), Johns Hopkins University, 2020), 4.

24 Willem H. Gravett, "Digital Neocolonialism: The Chinese Surveillance State in Africa," African Journal of International and Comparative Law 30, no. 1 (February 1, 2022): 41, https://doi.org/10.3366/ajicl.2022.0393.

25 Joe Parkinson, Nicholas Bariyo, and Josh Chin, "Huawei Technicians Helped African Governments Spy on Political Opponents," Wall Street Journal, August 14, 2019, https://www.wsj.com/articles/huawei-technicians-helped-african-governments-spy-on-political-opponents-11565793017.

26 Marczak et al., "Hide and Seek."

27 Chris Alden et al., "FOCAC at 21: Future Trajectories of China-Africa Relations" (LSE IDEAS, October 2021), https://www.lse.ac.uk/ideas/Assets/Documents/reports/LSE-IDEAS-FOCAC-at-21.pdf.

28 Mandira Bagwandeen, "Don't Blame China for the Rise of Digital Authoritarianism in Africa," Africa at LSE, September 9, 2021, https://blogs.lse.ac.uk/africaatlse/2021/09/09/dont-blame-china-for-rise-of-digital-authoritarianism-africa-surveillance-capitalism/.

29 Gagliardone, "Impact of Chinese Tech Provision."

30 Diego A. Martin, Jacob N. Shapiro, Michelle Nedashkovskaya, "Recent Trends in Online Influence Efforts" 18 no. 3 (2019): 23 https://www-jstor-org.uplib.idm.oclc.org/stable/26894680.

31 Gagliard Africa Center for Strategic Studies, "Mapping Disinformation in Africa," April 26, 2022, https://africacenter.org/spotlight/mapping-disinformation-in-africa/.

32 Martina Schwikowski, "Russia Targets Africa with Propaganda Machine," Deutsche Welle, November 29, 2022, https://www.dw.com/en/russia-targets-africa-with-propaganda-machine/a-63916836; Shannon Bond, "A Pro-Russian Social Media Campaign Is Trying to Influence Politics in Africa," NPR, February 1, 2023, https://www.npr.org/2023/02/01/1152899845/a-pro-russian-social-media-campaign-is-trying-to-influence-politics-in-africa.

33 Alex Hern and Luke Harding, "Russian-Led Troll Network Based in West Africa Uncovered," The Guardian, March 13, 2020, https://www.theguardian.com/technology/2020/mar/13/facebook-uncovers-russian-led-troll-network-based-in-west-africa.

34 Grigor Atanesian, "Russia in Africa: How Disinformation Operations Target the Continent," BBC News, February 1, 2023, https://www.bbc.com/news/world-africa-64451376.

35 Shelby Grossman, Daniel Bush, and Renée DiResta, "Evidence of Russia-Linked Influence Operations in Africa" (Stanford Internet Observatory, October 29, 2019), https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/29oct2019_sio_-_russia_linked_influence_operations_in_africa.final_.pdf.

36 CIPESA, "Digital Authoritarianism and Democratic Participation in Africa." Richard A. Wilson, "Digital Authoritarianism and the Global Assault on Human Rights," Human Rights Quarterly 44 no. 4 (2022): 709-712, https://muse-jhu-edu.uplib.idm.oclc.org/pub/1/article/868839/pdf.

37 CIPESA, "Digital Authoritarianism and Democratic Participation in Africa."

38 Privacy International, "For God and My President: State Surveillance in Uganda," October 2015, https://privacyinternational.org/sites/default/files/2017-12/Uganda_Report_1.pdf.

39 International Commission of Jurists, "Regulation of Communications Surveillance and Access to Internet in Selected African States," 2021, https://www.kas.de/documents/275350/0/Report-on-Regulation-of-Communications-Surveillance-and-Access-to-Internet-in-Selected-African-States.pdf/66dbd47d-4d7d-2779-a595-a34e9f93cfbb?t=1639140695434.

40 Lukman Adebisi Abdulrauf, "The Challenges for the Rule of Law Posed by the Increasing Use of Electronic Surveillance in Sub-Saharan Africa," African Human Rights Law Journal 18, no. 1 (2018): 367, https://doi.org/10.17159/1996-2096/2018/v18n1a17.

41 CIPESA, "Digital Authoritarianism and Democratic Participation in Africa."

42 Identification for Development, "Digital ID and the Data Protection Challenge: Practitioners Note" (World Bank, October 2019), https://documents1.worldbank.org/curated/en/508291571358375350/pdf/Digital-ID-and-the-Data-Protection-Challenge-Practitioners-Note.pdf.

43 African countries that have digital ID systems, data protection laws and functional DPAs include Angola, Benin, Botswana, Burkina Faso, Cape Verde, Mauritania, Niger, São Tomé & Príncipe, Nigeria, Senegal, Mali, Côte d'Ivoire, South Africa, Uganda, Kenya, Zimbabwe, Gabon, Rwanda, Ghana & Chad. Compare 40 above with UNCTAD, "Data Protection and Privacy Legislation Worldwide," accessed March 4, 2023, https://unctad.org/page/data-protection-and-privacy-legislation-worldwide.

44 Evelyn Mary Aswad, "Losing the Freedom to Be Human," Columbia Human Rights Law Review 52, no. 1 (2020): 361, https://hrlr.law.columbia.edu/files/2020/11/306_Aswad.pdf.

45 Tony Roberts and Tanja Bosch, "Case Study: Digital Citizenship or Digital Authoritarianism?," in Development Co-Operation Report 2021: Shaping a Just Digital Transformation (Paris: OECD Publishing, 2021), https://doi.org/10.1787/ce08832f-en.

Part II
Infrastructures of Repression and Resistance

In this section:

- An Infrastructural Approach to Digital Authoritarianism
- Panopticons and Closing Civic Space: The Building Blocks of Digital Authoritarianism
- Protecting Digital Infrastructures, Assets, and Users From Bad Actors: An Empowering Story of Jamiiforums From Tanzania

Berggruen Institute | Canada | UCLA Institute for Technology, Law & Policy | USC Dornsife Center on Science, Technology, and Public Life

An Infrastructural Approach to Digital Authoritarianism

Mike Ananny
University of Southern California*

I. Digital Authoritarianism as Infrastructural

Starting from a standard working definition of digital authoritarianism (DA) – "the use of information technology by authoritarian regimes to surveil, repress, and manipulate domestic and foreign populations"[1]–I want to suggest that the concept can more richly be seen as the unchecked creation and deployment of oppressive communication infrastructure. A focus on infrastructure offers a broader and more powerful way forward for counter-authoritarianism than the emphasis on tools and technologies that often dominates DA discussions.

I try to illustrate this claim with two stories, a definition of DA infrastructure, and a short tour of places where DA infrastructure appears today. I end by suggesting three ways that an infrastructural view of digital authoritarianism could drive new approaches to counter-authoritarianism.

II. Two Stories

In Shake Hands with the Devil,[2] his memoir as the force commander of the United Nations (UN) Assistance Mission during the 1993-94 Rwandan genocide, Canadian Major-General[3] Roméo Dallaire recounts how a chilling infrastructural calculus fueled an authoritarian slaughter. He had been pleading with the UN to shut down Rwanda's Radio Télévision Libre des Mille Collines (RTLM) station, a major communication channel that was fomenting violence, directing attacks, and officially sanctioning genocide. Dallaire determined that countless lives could be saved if RTLM were eliminated through some combination of "jamming, a direct air strike on the transmitter, or covert operations."[4] Only the United States military had the resources to do any of these three things but the Pentagon decided not to intervene against RTLM. It concluded that the cost ($8500 per hour for a signal-jamming aircraft) and legal uncertainties (blocking a country's radio waves would violate national sovereign) were too great to justify, even though its own research agreed with Dallaire's assessment in estimating that destroying the radio station would likely have saved 8,000-10,000 Rwandans. Dallaire later wrote that the "haunting image of killers with a machete in one hand and a radio in the other never leaves you... In some villages, radio was like the voice of God."[5]

Fast forward to 2017 and Myanmar's Facebook-fueled atrocities against the Rohingya. As the United Nations,[6] Amnesty International[7], and communication scholars[8] have documented, Facebook's (now Meta) quickly dominated almost all aspects of everyday life in Myanmar. This monopoly made it possible for state security forces to use the platform to coordinate attacks on the Rohingya quickly and effectively. As many people in

Myanmar observed at the time, Facebook is the Internet."[9] The US company had the power to shape and control the country's online life, with devastating impacts on the Rohingya people:

- "Meta's content-shaping algorithms were amplifying harmful content, including advocacy of hatred, in Myanmar as early as 2014"[10]
- The UN found Facebook's responses to authoritarian uses of its platform to be "slow and ineffective" and the company was unable to provide "country-specific data about the spread of hate speech on its platform"[11]
- Though Facebook chief Mark Zuckerberg testified in 2018 to the US Congress that Facebook would hire "dozens of more Burmese-language content reviewers, because hate speech is very language-specific" one investigation found that, at the time, the company "had only five Burmese language speakers to monitor and moderate content for Myanmar's 18 million Facebook users."[12]
- Facebook labeled four anti-government organizations "dangerous organizations," banning them from using the platform to organize counter-authoritarian actions[13]

In these stories of both Rwanda and Myanmar we see authoritarianism and genocide enacted not just by "information technologies" or through neutral "tools" but through a broader, more subtle, systematic, and messier set of infrastructures. These infrastructures play out in:

- legal theories of national sovereignty that respect radio waves;
- financial costs of deploying life-saving airplanes;
- bureaucratic calculations of military operating costs versus human lives;
- violence emerging from confluences of portable technologies like radios and machetes;
- ranking algorithms that make hate speech seem ubiquitous and popular;
- corporate timelines, definitions, data practices, and strategic priorities that hamper counter-authoritarian efforts;
- technology company CEOs placating regulators with polished admissions of failures to manage the complexities of their own systems, with promises to do better;
- and English-dominated technological cultures that leave many marginalized populations literally unheard.

Seen as such, authoritarianism is not about using tools but creating and deploying infrastructures. Though some studies suggest broader and more complexly intertwined socio-technical images of DA,[14] an infrastructural view is both a more accurate image of authoritarian power and a better framework for digital counter-authoritarianism.

III. Defining Infrastructure

Following long-standing Science and Technology Studies (STS) scholarship, I define infrastructures as largely hidden and taken-for-granted relationships among people, material, and language. They bring people, objects, and built environments together. They make some actions seem routine, expected, and "normal". They rely on a network of people and standards that are usually out of the public eye. And they often only become visible when they break down. Railroads, electricity grids, sewers, streetlights, air traffic control, undersea internet cables, and geosynchronous

satellites are all infrastructures that need vast and complex relationships among people, materials, machines, skills, and rules to operate successfully. Most people never give them a second thought – until a train derails, lights go out, A/C fails, flights are delayed, websites do not load, or algorithms make nonsensical recommendations. In a relatively short time, the internet, social media platforms, algorithmic systems, and artificial intelligence have all started behaving like infrastructures.[15]

Infrastructures are both utterly boring and deeply political. Indeed, part of their power comes from their ability to hide in plain sight and draw little attention while governing almost all aspects of life.

They depend upon seemingly neutral tests, terms, categories, and definitions that are often developed by bureaucracies and weakly connected to expert judgment or critical review, but that are created and deployed with seeming objectivity. An air traffic controller's "safe" separation of planes may be carefully calculated, publicly regulated, and reflect years of revision in light of changes in aircraft technology, near-misses, and crash investigations. In contrast, many of the seemingly objective concepts, language, and guidelines driving social media platforms (like "hate speech", "political ad", "community standard", "trust and safety") are new, privately developed, and hotly contested.[16] These words—and the politics underlying them—are what make infrastructural breakdowns of social media so hard to trace and understand as powerful. It is easy to see when a train goes off the tracks, a cable breaks, or a satellite drifts out of orbit, but it is much harder to agree that a recommendation algorithm is "right", a machine learning systems is "good enough", or that a social media platform is "safe."

The world is made of infrastructures with genealogies—layers of history, skill, judgment, assumption, technique, language, material, mistakes, revision, and contestation that may seem boring and neutral, but are actually deeply political.

If you accept an infrastructural view of DA then new approaches to counter-authoritarian come into focus. All of the ways that digital infrastructures are subtly made, politically powerful, and constantly unfolding become starting points for reform and resistance. If DA is infrastructural, where exactly are its relationships? And how do these relationships offer new ways to counter authoritarianism?

IV. Infrastructural Approaches to Digital Authoritarianism

In his 2020 Massey Lectures, University of Toronto cybersecurity scholar Ron Deibert argues that although digital technology platforms enable

> a new type of 'digital authoritarianism' to flourish, it would be a mistake to conclude that the effects of social media are limited to a group of bad actors 'over there.' In fact, the most disturbing dynamics are playing themselves out within nominally liberal democratic countries[17]

One way to see how digital authoritarianism is not just "over there" is to examine the digital communication infrastructures in ostensibly stable and sophisticated Western democracies.[18] But, as STS scholars show, well-working infrastructures like domestic DA infrastructures are usually invisible. They can become visible in dramatic moments—like the so-called Canadian "Freedom Convoys" of 2022 or the January 6th attacks on the US Congress—but they are otherwise hard to see. Scholars, activists, and policymakers can see infrastructures and their invisible power by learning to do what STS scholars Bowker and Star call "infrastructural inversion."

This method involves looking

> closely at technologies and arrangements that, by design and by habit, tend to fade into the woodwork... recognizing the depths of technical networks and standards, on the one hand, and the real work of politics and knowledge production on the other... [giving] them causal prominence in many areas usually attributed to heroic actors, social movements, or cultural mores.[19]

Though there is not space here to fully explore applying infrastructural inversion to DA, I want to use the approach to show at least five ways that DA appears within largely invisible, and seemingly apolitical domestic digital communication infrastructures:

- The Charismatic Genius
- Strategic Simplifications
- Endless Iteration
- Tamed Ideals of the Public
- Synthetic Media

I briefly discuss each of these dynamics, aiming to show how seemingly innocuous and mundane infrastructural relationships hide domestic digital authoritarianism in plain sight.

1. The Charismatic Genius

As Morgan Ames shows in her critical study of the One Laptop Per Child program,[20] as Rob Reich argues in his history of US philanthropy,[21] and as Julia Sonnevend shows in her examination of Victor Orban's "charm",[22] many Western, domestic images of progress depend upon the myths of charismatic genius and the largess of individuals (mostly men) with superhuman intelligence, inexhaustible personal fortunes, and visions of new worlds.[23] The idea that geniuses can save societies is a subtle form of authoritarian appearing today in things like fawning profiles of Elon Musk's inventiveness;[24] hopes that Jeff Bezos will save journalism[25] or that Bill Gates will end climate change;[26] Mark Zuckerberg's (discontinued) annual "personal challenges" of individual and societal transformation;[27] the interstellar colonialism of Richard Branson's once celebrated "Virgin Orbit" venture;[28] Peter Thiel's "seasteading" on artificially constructed, utopic floating islands liberated from politicians;[29] or Twitter adjusting its content moderation policies on the fly, to accommodate Trump tweets that would otherwise have violated its community guidelines.[30] The charismatic genius says that he can fix this world or invent an entirely new one.

The largely Western myth of a person smart, powerful, and wealthy enough able to create change through infrastructure is a kind of invisible, background ideology that leaves many of the assumptions driving infrastructures unseen and unchecked. The charismatic genius may spend his money with good intentions and spur activity that seems to liberate us from bureaucracy. But he is also a dangerous and authoritarian figure because of his assumption that complex public problems can be solved through the singular vision of wealthy individuals. He makes politicians and publics alike beholden to his creativity, expertise, and infrastructural interests – a subtle but powerful mix of charisma, hope, celebrity, charm, and authoritarianism.

2. Strategic Simplifications

Similarly, in C.W. Anderson's dissection of tech manifestos,[31] David Karpf's archival analyses of WIRED magazine rhetoric,[32] and Fred Turner's history of democratic technologies,[33] we can see how charismatic geniuses and techno-utopian cultures boil down complex and nuanced ideas—like free speech—into soundbites, marketing, and zealotry. At Georgetown University, Mark Zuckerberg asserted a theory of free speech that just happened to align with Facebook's

business model;[34] in a now deleted blog post Twitter co-founder Jack Dorsey offered both an apology for the platform's decline under Elon Musk and a new theory of free speech infrastructure aligned with his latest venture;[35] and Elon Musk recently collapsed Twitter's complex recommendation infrastructure into a single design that highly ranked his own tweets.[36]

Charismatic geniuses often craft strategic simplifications in order to see hard problems in their own terms, as products and policies underpinning and advancing the infrastructures that they control. These simplifications are enticing and comforting with an almost religious promise of betterment, but they need faith, control, and bounded definitions. Simplifications promise focus, stability, and efficiency. They recast complexities within and across massively scaled communication platforms, climate crises, pandemics, political unrest, artificial intelligence—what Candace Callison calls a "syndemic" view of contemporary life[37]--into actions and solutions that need authoritative infrastructures.

3. Endless Iteration

As Neff and Stark argue in "Permanently Beta"[38] and Hapt does in "Facebook Futures",[39] technology companies are always promising that better days are almost here. They promise that they are continually improving, that they are listening to customers and regulators, that they are investing in better products and policies, and that they can be trusted to improve the lives of people and the health of society if they are just given enough time, self-direction, and freedom to fail on their own terms. They apologize, sympathize with users, minimize fallout through crisis communication techniques, cite the complexity and unknowability of their own systems as reasons for errors, and try to reposition corporate failings as symptoms of larger social and cultural forces beyond their control.[40] The charismatic geniuses who "move fast and break things" promise that more iteration will lead to neater simplifications, stronger problem definitions, and better solutions.

Users, citizens, regulators, politicians and activists must work within these cycles of iteration. They have to study each cycle of shock and exception[41] for clues about how it is like or unlike earlier innovations. They must ask whether they can withstand more experimentation before fleeing for an alternative infrastructure (if one exists) and try to see how failures are unevenly distributed to the most vulnerable in society. They need to ask hard questions about whether they really have the power to counteract endless iteration, and what the costs would be of rejecting infrastructures that are not getting better fast enough. The ritualized apologies and confessions of the charismatic geniuses become ways of stalling, controlling timelines, and crafting a seemingly reasonable commonsense about how technologies are genies that cannot be put back in the bottle, just worked with and ameliorated. The largely invisible and normalized narrative of endless infrastructural iteration seems to be the only rational way forward in the face of unavoidable, inevitable harms that might, eventually, be lessened. This loop is a subtle form of infrastructural authoritarianism.

4. Tamed Ideals of the Public

DA infrastructure can also subtly and gradually control and compress the meaning of "public," making public life less easy to imagine and realize as anything other than what authoritarians prefer.

Though the word has a long, fraught, and nuanced history—defined as everything from state performances[42] and media institutions[43] to shared social consequences,[44] resistance movements,[45] and data cultures[46]— contemporary creators of digital infrastructures routinely

manipulate the concept of publicness. Twitter maintains a list of "public-interest exceptions" to its policies that includes its own corporate definition of "public."[47] The committee investigating the January 6th congressional attacks found that Twitter failed to follow these exceptions, "to avoid penalizing conservatives, particularly then-president Trump."[48] Since Elon Musk purchased Twitter, the platform has banned journalists without due process and hosted more climate misinformation and hate speech than ever before – all under the guise of creating a more open forum for public debate.[49] TikTok regularly uses an amplification technique called "heating"—usually reserved by social media platforms for content like public health warnings or election information—to instead "court influencers and brands, enticing them into partnerships by inflating their videos' view count," blurring the lines between public-interest and business-driven amplification.[50] And even after Musk's authoritarian actions after acquiring Twitter—banning journalists, hosting misinformation, leaking private messages—many cities still refuse to stop using Twitter for civic communication, claiming that they have few other good options and that they are powerless to counter Musk's authoritarian changes. Even if a city's constituents wanted to avoid using Twitter as a source of public information, city administrators seem unwilling to reject Twitter and it remains a key platform for public information and discourse.[51] Even more insidious, the platform can become the only place to express outrage about the power of the platform and call for its redesign, as was the case when suicide prevention advocates found themselves needing to use Twitter to demand that Musk reinstate Twitter's suicide-prevention tools.[52]

While there is no agreed upon meaning of "public" among scholars, activists, and policymakers, when authoritarian infrastructures have the power to define the public interest, unilaterally make public interest exemptions, and capture the spaces where public officials communicate with constituents, public life becomes smaller, narrower, easier to manipulate. The meaning and realization of "public" becomes beholden to the geniuses, simplifications, and iterations of DA infrastructures.

5. Synthetic Media

Finally, as recent scholarship and journalism show, one of the newest and potentially most powerful forms of DA infrastructure is systems that fabricate media entirely.

Using vast, often proprietary, and under-studied datasets and "large language models,"[53] actors of various kinds (corporations, states, artists, journalists, activists) can now use machine learning to create text, audio, imagery, video of all kinds that look like original, human-created media but that are actually generated by machines. Some scholars go so far as to suggest that artificial intelligence systems are state actors.[54] Unlike earlier eras of media criticism that traced the people, organizations, and interests behind news narratives and public relations campaigns,[55] this machine-generated content emerges from a poorly understood mix of computational systems that can, by design, include "hallucinated," fictional information. It is not only hard to know whether such media are truthful or the extent to which people were involved in their creation, it can be virtually impossible to see and trace all the statistical systems driving them – a kind of oppression-by-automation that resists traditional forms of critique and resistance. Variously called "synthetic media" or "generative artificial intelligence" such expressions can range from innocuous and entertaining to nefarious and oppressive – e.g., an AI-created children's story describing a complex scientific phenomenon, disaster simulations that spur preparedness, or

fabricated stories that ruin reputations, spur investigations, or strategically distract news audiences.[56]

While the power of synthetic media is still unclear, it already has the hallmarks of DA infrastructure – largely invisible data, specialized expertise, vast and complex systems, seemingly boring operations that can cause spectacular failures, uncertain outcomes that are hard to anticipate and prevent

V. Ways Forward

Considering these five dimensions together, a picture emerges of digital authoritarianism that is more than just the use of neutral tools by oppressive actors. An infrastructural view of DA shows how authoritarianism does not have to be overt oppression; rather, as digital authoritarianism scholar Ron Deibert argues, oppression can also be "confusion, ignorance, prejudice, and chaos."[57] An infrastructural model of DA shows how subtle forms of authoritarianism can slowly emerge from social, cultural, and technological forces that are often invisible, taken for granted, and resting upon ideologies of genius, simplification, iteration, tamed public, and hallucinated realities.

What is to be done? I suggest three recommendations.

First, as STS scholar Leigh Star urged many years ago, fund and celebrate the study of "boring things."[58] While there is tragically no shortage of headline-grabbing examples of violent and oppressive authoritarianism, we should prioritize the study of seemingly uninteresting or innocuous infrastructures in two ways: (a) look for infrastructural forces in the obvious examples of authoritarian oppression; and (b) go the other direction to ask what forms of authoritarian oppression infrastructural forces could create.

Second, demystify infrastructures. Resist the explanations, excuses, and apologies offered by the self-defined geniuses. Question who benefits from infrastructures that are too big and complex to be traced and held publicly accountable. Call out the power that comes from creating and deploying opaque systems. See how infrastructural failures often exacerbate injustices and inequitably distribute harm to those already lacking social and economic power. Demand access to the cultures, communities, and closed rooms that create complex infrastructures; do not mistake access to server data with cultural insight and public accountability. Identify what makes DA infrastructures both powerful and vulnerable, and then attack their infrastructural weaknesses in ways that lessen their power.

Finally, imagine and work to realize alternative infrastructures that can counter digital authoritarianism (e.g., end-to-end encryption, federated platforms like Mastodon, business models that reject surveillance capitalism). Simultaneously, understand the trade-offs that new, alternative, seemingly liberatory counter-authoritarian infrastructures always entail.[59] Be able to articulate the public standards that counter-authoritarian systems should meet and do not confuse the community dynamics of counter-authoritarian hackers with the public interests of diverse, infrastructural constituents.[60]

By centering an infrastructural model of digital authoritarianism, I do not mean to minimize the difficulty, necessity, and power of shutting down a tool, blocking an actor, or using a more obvious form of technological force. Rather, by seeing digital authoritarianism as infrastructure we may see new, subtle, powerful, and currently invisible ways to counter oppression and build public life.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 "Digital Human Rights Need a Single Home in U.S. Government," Foreign Policy, updated March 14, 2022, 2022, https://foreignpolicy.com/2022/03/14/digital-authoritarianism-tech-human-rights/.

2 Roméo Dallaire, Shake Hands with the Devil: The Failure of Humanity in Rwanda (Toronto, ON: Carroll & Graf, 2005).

3 Dallaire was subsequently promoted to Lieutenant-General and later served as a Canadian Senator.

4 Dallaire, Shake Hands with the Devil: The Failure of Humanity in Rwanda, 375.

5 Roméo Dallaire, "The Media Dichotomy," in The Media and the Rwanda Genocide, ed. Allan Thompson (Pluto Press, 2007), 12, 16. See also Alison Des Forges, "Call to Genocide: Radio in Rwanda, 1994," in The Media and the Rwanda Genocide, ed. Allan Thompson (Pluto Press, 2007).

6 United Nations, Report of the Detailed Findings of the Independent International FactFinding Mission in Myanmar (September 28, 2018 2018), https://www.ohchr.org/sites/default/files/Documents/HRBodies/HRCouncil/FFM-Myanmar/A_HRC_39_64.pdf.

7 Amnesty International, Myanmar: The social atrocity: Meta and the right to remedy for the Rohingya (September 29, 2022 2022), https://www.amnesty.org/en/documents/ASA16/5933/2022/en/.

8 For example, see Jeffrey Sablosky, "Dangerous organizations: Facebook's content moderation decisions and ethnic visibility in Myanmar," Media, Culture & Society 43, no. 6 (2021), https://doi.org/10.1177/0163443720987751; Jenifer Whitten-Woodring et al., "Poison If You Don't Know How to Use It: Facebook, Democracy, and Human Rights in Myanmar," The International Journal of Press/Politics 25, no. 3 (2020), https://doi.org/10.1177/1940161220919666.

9 United Nations, Report of the Detailed Findings of the Independent International FactFinding Mission in Myanmar, 14.

10 Amnesty International, Myanmar: The social atrocity: Meta and the right to remedy for the Rohingya, 46.

11 United Nations, Report of the Detailed Findings of the Independent International FactFinding Mission in Myanmar, 14.

12 Amnesty International, Myanmar: The social atrocity: Meta and the right to remedy for the Rohingya, 34. Note that Dallaire reports a similar struggle to monitor Rwandan genocide communications in local languages: "We had so little capacity to monitor broadcasts, particularly those in the local language, Kinyarwanda. For a long time, we didn't notice the difference in tone between RTLM broadcasts in French and those in Kinyarwanda. We missed this vital early-warning sign of what was to come because, in effect, we weren't listening properly to local media and what it was telling people." Dallaire, "The Media Dichotomy," 16-17.

13 Sablosky, "Dangerous organizations: Facebook's content moderation decisions and ethnic visibility in Myanmar."

14 For example, see "Exporting digital authoritarianism: The Russian and Chinese models," Brookings Institute - Exporting Digital Authoritarianism, updated August, 2019, 2019, https://www.brookings.edu/wp-content/uploads/2019/08/FP_20190827_digital_authoritarianism_polyakova_meserole.pdf; "Authoritarianism Has Been Reinvented for the Digital Age," Center for International Governance Innovation, updated July 9, 2021, 2021, https://www.cigionline.org/articles/authoritarianism-has-been-reinvented-for-the-digital-age/; Ron Deibert, "Authoritarianism Goes Global: Cyberspace Under Siege," Journal of Democracy 26, no. 3 (2015), https://doi.org/10.1353/jod.2015.0051; Ron Deibert, "The Road to Digital Unfreedom: Three Painful Truths About Social Media," Journal of Democracy 30, no. 1 (2019), https://doi.org/10.1353/jod.2019.0002; Erol Yayboke and Samuel Brannen, Promote and Build: A Strategic Approach to Digital Authoritarianism, Center for Strategic and International Studies (2020), https://www.csis.org/analysis/promote-and-build-strategic-approach-digital-authoritarianism; Adrian Shahbaz, The Rise of Digital Authoritarianism, Freedom House (2018), https://freedomhouse.org/report/freedom-net/2018/rise-digital-authoritarianism.

15 The STS literature on infrastructure is too vast to thoughtfully summarize here, but excellent work on the concept includes: Paul Edwards, "Infrastructure and modernity: Force, time, and social organization in the history of sociotechnical systems," in Modernity and technology, ed. T.J. Misa, P. Brey, and A. Feenberg (Cambridge, MA: The MIT Press, 2003); Paul N. Edwards et al., "Introduction: An agenda for infrastructure studies," Journal of the Association for Information Systems 10, no. 5 (2009); Brian Larkin, "The politics and poetics of infrastructure," Annual Review of Anthropology 42 (2013), https://doi.org/10.1146/annurev-anthro-092412-155522; Lisa Parks and Nicole Starosielski, eds., Signal traffic: Critical studies of media infrastructure (Chicago, IL: University of Illinois Press, 2015); Jean-Christophe Plantin and Aswin Punathambekar, "Digital media infrastructures: pipes, platforms, and politics," Media, Culture & Society 41, no. 2 (2019), https://doi.org/10.1177/0163443718818376, https://doi.org/10.1177/0163443718818376; Susan Leigh Star, "Ethnography of infrastructure," American Behavioral Scientist 43, no. 3 (1999); Nicole Starosielski, The Undersea Network (Durham, NC: Duke University Press, 2015); Antina Von Schnitzler, Democracy's infrastructure: Techno-politics and protest after apartheid (Princeton, NJ: Princeton, 2016).

16 For a discussion of the political power that categories have to define digital infrastructures see Mike Ananny, "Making up political people: How social media create the ideals, definitions, and probabilities of political speech," Georgetown Law Technology Review 4, no. 2 (2020).

17 Ron Deibert, Reset: Reclaiming the Internet for Civil Society (Toronto, Canada: Anansi, 2020), np.

18 When reports on Digital Authoritarianism do discuss infrastructure, it is largely in terms of the dangers of Russian and Chinese digital infrastructures, not domestic Western infrastructures. E.g., see Polyakova and Meserole, "Exporting digital authoritarianism: The Russian and Chinese models."; Yayboke and Brannen, Promote and Build: A Strategic Approach to Digital Authoritarianism; Shahbaz, The Rise of Digital Authoritarianism.

19 Geof C. Bowker and Susan Leigh Star, Sorting things out: Classification and its consequences (Cambridge, MA: The MIT Press, 1999), 34.MA: The MIT Press, 1999

20 Morgan G Ames, The Charisma Machine: The Life, Death, and Legacy of One Laptop per Child (Cambridge, MA: MIT Press, 2019).

21 Rob Reich, Just Giving (Princeton, NJ: Princeton University Press, 2018).

22 Julia Sonnevend, "Charm offensive: mediatized country image transformations in international relations," Information, Communication & Society 22, no. 5 (2019/04/16 2019), https://doi.org/10.1080/1369118X.2019.1568516, https://doi.org/10.1080/1369118X.2019.1568516.

23 For an in-depth and critical history of the tensions behind tech-driven homesteading, escapism, and utopic creation see Fred Turner, From counterculture to cyberculture: Stewart Brand, the Whole Earth Network, and the rise of digital utopianism (Chicago, IL: University of Chicago Press, 2006).

24 Kenan Malik, "Beware self-made 'genius' entrepreneurs promising the earth. Just look at Elon Musk," The Guardian, November 20, 2022 2022, https://www.theguardian.com/commentisfree/2022/nov/20/beware-self-made-genius-entrepreneurs-promising-earth-just-look-at-elon-musk; Nellie Bowles, "Tesla Owners Try to Make Sense of Elon Musk's 'Red Pill' Moment," The New York Times, May 19, 2020 2020, https://www.nytimes.com/2020/05/19/technology/elon-musk-tesla-red-pill.html.

25 Stephanie Denning, "Why Jeff Bezos Bought The Washington Post," Forbes, September 19, 2018 2018, https://www.forbes.com/sites/stephaniedenning/2018/09/19/why-jeff-bezos-bought-the-washington-post/.

26 Bill Gates, How to Avoid a Climate Disaster (New York, NY: Penguin Random House, 2021).

27 Kurt Wagner, "Mark Zuckerberg Ends His Annual Challenge Tradition in Favor of a 'Longer-Term Focus'," Fortune, January 9, 2020 2020, https://fortune.com/2020/01/09/mark-zuckerberg-annual-personal-challenge/.

28 Alistair MacDonald, "Virgin Orbit, Richard Branson's Satellite-Launch Venture, Has Fallen Fast," Wall Street Journal, March 18 2023, https://www.wsj.com/articles/virgin-orbit-richard-bransons-satellite-launch-venture-has-fallen-fast-29cc8ea.

29 Joe Quirk, Seasteading (New York, NY: Simon Schuster, 2017).

30 Kari Paul, "Republicans grill ex-Twitter executives over handling of Hunter Biden story," The Guardian, February 8, 2023 2023, https://www.theguardian.com/us-news/2023/feb/08/ex-twitter-execs-to-testify-in-congress-on-handling-of-hunter-biden-laptop-reporting.

31 "Dissecting Tech Manifestos," Tech Policy Press, 2022, https://techpolicy.press/captivate-podcast/dissecting-tech-manifestos/.

32 "25 Years of WIRED Predictions: Why the Future Never Arrives," WIRED, 2018, https://www.wired.com/story/wired25-david-karpf-issues-tech-predictions/.

33 Fred Turner, "Machine politics: The rise of the internet and a new age of authoritarianism," Harper's Magazine 338 (2019).

34 "Mark Zuckerberg Addresses Students at Georgetown Event Kicking Off New Series," 2019, https://www.georgetown.edu/news/mark-zuckerberg-to-host-conversation-at-georgetown-on-free-expression/.

35 "a native internet protocol for social media," Revue / Wayback Machine, 2022, https://web.archive.org/web/20230117185515/https://www.getrevue.co/profile/jackjack/issues/a-native-internet-protocol-for-social-media-1503112.

36 Kari Paul, "Elon Musk reportedly forced Twitter algorithm to boost his tweets after Super Bowl flop," The Guardian, February 15, 2023 2023, https://www.theguardian.com/technology/2023/feb/15/elon-musk-changes-twitter-algorithm-super-bowl-slump-report.

37 Candis Callison, What COVID-19 and climate change teach us about "syndemics", Institute for Research on Public Policy (Ottawa, Canada, 2021), https://policyoptions.irpp.org/magazines/march-2021/what-covid-19-and-climate-change-teach-us-about-syndemics/.

38 Gina Neff and David Stark, "Permanently beta: Responsive organization in the internet era," in Society online: The internet in context, ed. Philip Howard and Steven Jones (Thousand Oaks, CA: Sage, 2004).

39 Joachim Haupt, "Facebook futures: Mark Zuckerberg's discursive construction of a better world," New Media & Society 23, no. 2 (2021).

40 See Kimberly Hall, "Public Penitence: Facebook and the Performance of Apology," Social Media + Society 6, no. 2 (2020), https://doi.org/10.1177/2056305120907945; Tero Karppi and David B Nieborg, "Facebook confessions: Corporate abdication and Silicon Valley dystopianism," New Media & Society 23, no. 9 (2021), https://doi.org/10.1177/1461444820933549.

41 Mike Ananny and Tarleton Gillespie, “Public Platforms: Beyond the Cycle of Shocks and Exceptions” (The Platform Society, Oxford, UK, Oxford Internet Institute, 2016).

42 John Durham Peters, “Historical tensions in the concept of public opinion,” in Public Opinion and the Communication of Consent, ed. Theodore L. Glasser and Charles T. Salmon (New York: The Guildford Press, 1995).

43 Jurgen Habermas, Between facts and norms: Contributions to a discourse theory of law and democracy (Cambridge, MA: MIT Press, 1996). .

44 John Dewey, The public and its problems (New York: Swallow Press, 1954).

45 Catherine R. Squires, “Rethinking the black public sphere: An alternative vocabulary for multiple public spheres,” Communication Theory 12, no. 4 (2002), https://doi.org/10.1111/j.1468-2885.2002.tb00278.x.

46 Slavko Splichal, Datafication of Public Opinion and the Public Sphere (London, UK: Anthem Press, 2022).

47 “Defining public interest on Twitter,” Twitter Safety, updated June 27, 2019, 2019, accessed May 2, 2020, https://blog.twitter.com/en_us/topics/company/2019/publicinterest.html.

48 Cat Zakrzewski, Cristiano Lima, and Drew Harwell, “What the Jan. 6 probe found out about social media, but didn’t report,” The Washington Post, January 17, 2023 2023, np, https://www.washingtonpost.com/technology/2023/01/17/jan6-committee-report-social-media/.

49 Sheera Frenkel and Kate Conger, “Hate Speech’s Rise on Twitter Is Unprecedented, Researchers Find,” The New York Times, December 2, 2022 2022, https://www.nytimes.com/2022/12/02/technology/twitter-hate-speech.html; Carl Miller et al., Antisemitism on Twitter Before and After Elon Musk’s Acquisition, Institute for Strategic Dialogue (London, UK, March 20, 2023 2023), https://www.isdglobal.org/isd-publications/antisemitism-on-twitter-before-and-after-elon-musks-acquisition/.

50 Emily Baker-White, “TikTok’s Secret ‘Heating’ Button Can Make Anyone Go Viral,” Forbes, January 20, 2023 2023, https://www.forbes.com/sites/emilybaker-white/2023/01/20/tiktoks-secret-heating-button-can-make-anyone-go-viral/.

51 Sarah Holder, “Are Cities Too Reliant on Twitter?,” Bloomberg City Lab Government (New York, NY), January 13 2023, https://www.bloomberg.com/news/articles/2023-01-13/despite-chaos-cities-are-sticking-with-twitter.

52 Edward Helmore, “Twitter restores suicide-prevention hotline feature after outcry,” The Guardian, December 24 2022, https://www.theguardian.com/technology/2022/dec/24/elon-musk-twitter-feature-removal-reports.

53 Alexandra Sasha Luccioni et al., “A Framework for Deprecating Datasets: Standardizing Documentation, Identification, and Communication” (2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, Association for Computing Machinery, 2022).

54 Kate Crawford and Jason Schultz, “AI systems as state actors,” Columbia Law Review 119, no. 7 (2019). For earlier scholarship making a similar claim that news organizations are also state actors, see Timothy E. Cook, Governing with the news (Chicago, IL: University of Chicago Press, 1998).

55 E.g., see William A. Gamson et al., “Media images and the social construction of reality,” Annual Review of Sociology 18 (1992).

56 For more on these dynamics and examples, see Gary King, Jennifer Pan, and Margaret E. Roberts, “How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument,” American Political Science Review 111, no. 3 (2017), https://doi.org/10.1017/S0003055417000144; Tiffany Hsu, “As Deepfakes Flourish, Countries Struggle With Response,” The New York Times, January 22, 2023 2023, https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html; Europol Innovation Lab, Facing reality? Law enforcement and the challenge of deepfakes, Europol (2022), https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf; Josh A. Goldstein et al., Forecasting potential misuses of language models for disinformation campaigns—and how to reduce risk, Stanford Internet Observatory (2023), https://cyber.fsi.stanford.edu/io/news/forecasting-potential-misuses-language-models-disinformation-campaigns-and-how-reduce-risk; Nathan E. Sanders and Bruce Schneier, “How ChatGPT Hijacks Democracy,” The New York Times, January 15, 2023 2023, https://www.nytimes.com/2023/01/15/opinion/ai-chatgpt-lobbying-democracy.html.

57 Deibert, “The Road to Digital Unfreedom: Three Painful Truths About Social Media,” 31.

58 Star, “Ethnography of infrastructure.”

59 As Dallaire wrote, his forces in Rwanda “desperately needed its own media outlet, its own radio station. We were unarmed in the media war that was going on and had virtually no capacity to explain ourselves to the local community to whom radio was so important… People saw a white vehicle with a blue flag going by at 70 kilometers an hour, but many Rwandans had no idea why we were there.” Dallaire, “The Media Dichotomy,” 17.

60 For an excellent discussion of the tensions between hacker community ethics and political theories of the public see Gabriella Coleman, Coding freedom: The ethics and aesthetics of hacking (Princeton, NJ: Princeton University Press, 2013).

Panopticons and Closing Civic Space: The Building Blocks of Digital Authoritarianism

Courtney C. Radsch
UCLA*

I. Democratic Values at Risk

The combination of generative AI, ever increasing computational power, and insatiable datafication provides authoritarian governments with the tools to monitor and control their citizens more effectively than ever before, leading to a further erosion of freedom and autonomy. Yet as this essay highlights, there is a perilously thin line separating digital authoritarianism from digital governance and public safety initiatives more broadly.

Digital authoritarianism refers to the use of technology and digital platforms by those in power to control, monitor, and manipulate public opinion and information and restrict civic space. It often focuses on tactics -- such as censorship, surveillance, propaganda, and manipulation of online content – to the exclusion of the underlying sociotechnical infrastructure and the capabilities that are being built. This is a shortcoming that must be addressed as advanced technologies and algorithmic systems become increasingly pervasive and creating new potentialities to turbocharge digital authoritarianism. The rapid pace of AI advances and their deployment throughout all facets of daily life exacerbate the potential to deploy prophetic and predictive manipulation, especially when the underlying economic and political incentives are not aligned with democratic or liberal emancipatory values.

This essay examines how the fabric of digital authoritarianism is being woven through monitoring, datafication and AI, shrouding the public sphere and civic space in surveillance and control and challenging the notion that there is meaningful distinction to be made between digital authoritarianism and digital democracy, meaning that it is only the political designation that distinguishes the two. I illustrate this in three parts, and consider the implications for journalism. The press has been considered a fundamental pillar of democratic systems, and a free press correlates with good governance and less corruption, but digital authoritarianism imperils journalism and risks rendering the press obsolete. Part II focuses on understanding the building blocks of digital authoritarianism, which are often indistinguishable from the building blocks of any other type of political system in the age of AI. Part III examines how the misalignment of political and economic values perverts the public sphere through disinformation, propaganda, and harassment and what this could mean for the future of journalism and independent media. Part IV analyzes how national ID cards, social credit systems, and other forms of surveillance lead to closing civic space, human rights abuses, and the impossibility of anonymity. The lack of anonymity means, for example, that journalists can no longer offer sources protection and whistleblowers will face far steeper penalties.

II. The Building Blocks: Data, Surveillance, AI, and Computing Power

Advances in engineering and manufacturing have permitted hereto untold types of information and signals to be turned into data. Extraordinary advances in computational power coupled with decreasing computational costs enable governments and the private sector to process and analyze ever larger amounts of data.[1] The application of complex math and machine learning approaches like reinforcement learning, adversarial learning, and neural networks to make sense of all this data produces new insights, learning, and predictive possibilities. But these also enable governments to more effectively track and monitor citizens and contribute to closing civic space worldwide.

The root of the datafication problem is the creation of certain types of data in the first place. Then its external storage and ownership model, since most data is collected, stored, and "owned" by third parties. This enables its commodification, aggregation, and integration with other systems, which in turn fuels machine learning and AI systems. "The more the computing power, the faster we can feed the data to train the AI system, resulting in a shorter span for the AI to reach near-perfection," as the CEO of a fintech institution put it.[2]

Much of this datafication can be beneficial to our daily lives, used to improve health, public safety, the environment, and the like. But the relentless collection, storage, and analysis of vast amounts of personal and collective data in public, private, and intimate spaces[3] also lays the groundwork for digital authoritarianism through improved decision-making and predictive algorithms that can then be used to deter and prevent undesirable speech or behavior.

Even with appropriate safeguards in place, they can too easily be lifted or abrogated amid shifting political winds or economic developments.

Surveillance, monitoring, and datafication are pervasive and inescapable in many countries and communities.[4] More and more data are being created and collected in civic spaces, from the digital public sphere to the physical public spaces where daily life was once lived in relative anonymity. From ubiquitous CCTV cameras to the plethora of sensorial and surveillance technologies that pervade modern life, the evolution toward "smart" cities and enhancing public services through "smart" national identify mechanisms, anonymity is an increasingly ephemeral value in digital age.

So, too, is privacy. Through our engagement with IoT devices, biometric monitoring, and DNA we invite our most intimate selves to be turned into data points and invite data collection through surveillance into the traditionally private space of the home. Data about our movements, our networks, our habits, and our preferences are collected through our engagement online, our devices, our purchases, and increasingly our simple existence in modern life.

The soma of convenience and innovation have created myriad modern-day panopticons that form the backbone of networked AI surveillance infrastructure, which is becoming an inescapable part of modernity.[5]

Our online expression is collected, analyzed, and deployed to create artificial intelligence systems and build and train algorithms that govern the public sphere, make decisions about risk and freedom, and shape the future. This can be seen in the realm of the public sphere, where content moderation and generative AI enable influence operations to flourish and risk turning journalism into an obsolete practice of a prior age.

And with the deployment of algorithmic decision-making in an increasing number of domains, from housing and health to criminal justice and law enforcement.

Algorithms increase the impact of the past on the present and future.[6] They reinforce structural and historic inequities and biases, hegemonic perspectives, and reproduce representational harms. Natural language processing (NLP), for example, learns from the vast swaths of digital texts that have been created by human and non-humans alike; news media, Wikipedia, Reddit, Twitter, and other social media platforms are important sources of training data for NLP systems,[7] which means that the propaganda of state media, the online harassment of social media, and the digital detritus of state-aligned influence operations are incorporated into the basis for learning in these systems.

These algorithmic and machine learning advances enable predictive manipulation, which attempts to anticipate and influence people's behavior through targeted messages or 'nudges.' This type of technology currently uses both algorithmic-based methods to capture data, such as web histories, purchase histories, and profiles on social media networks, as well as psychological methods that seek to manipulate users into thinking certain things or taking certain actions based on predictions about what they will respond to. But with more and better data aggregation and increased computational power, the power of predictive manipulation is likely to increase exponentially and be used in a wider range of domains.

In addition to existing behavioral surveillance systems that can now be deployed, incredible computation power can be used to learn about how diverse types of data and information are connected. Public surveillance systems, geotracking, identity cards, financial transactions and social network information all become fodder for artificial intelligence. AI that improves facial recognition, biometric surveillance, computer vision and natural language processing powers generative systems creates more and better data that can be reincorporated and reintegrated into AI systems in a self-reinforcing cycle that gives those systems greater potential economic and political power.

III. Information Operations, Generative News, and the Digital Public Square

Information operations targeting domestic and foreign populations, deploying spyware and leveraging influence operations to target the press and manipulate public opinion, attitudes and actions in order to support the political goals of the state's leaders or influence elections are common aspects of digital authoritarianism.[8] Whether loosely coordinated or tightly choreographed campaigns, they leverage state and party resources, including state and pro-government media, as well as public relations firms and armies of digital volunteers. Generative AI like ChatGPT and StabilityAI drastically reduce the time, money, and skill needed to drown the public sphere in propaganda and conduct influence and harassment campaigns.

An exponential increase in the budgets, personnel, and attention devoted by states, governments, and political parties to information operations has occurred amid a decline economic stability of independent media. While China and Russia have been among the most prolific users of online propaganda, they are far from alone. Political parties or government agencies in at least 70 countries deployed social media manipulation campaigns to influence domestic public opinion, and at least 45 countries deployed computational propaganda campaigns during elections, though this undoubtedly increases each year.[9] This means they used 'political bots' to amplify hate speech, manipulate

content, or get legitimate content removed; illegally harvested data and deployed micro-targeting; or mobilized cyber troops to bully or deter free speech and political opposition in campaigns explicitly or implicitly sanctioned by the state. Influence operations are increasingly common during elections because as they become more pervasive they become cheaper and it becomes harder to win without them.[10] This is part of what is fueling the deterioration of the public sphere worldwide. According to V-Dem, last year a record 35 counties experienced serious deteriorations in freedom of expression propelled by government, with civil society repression and media censorship worsening in more than 20 countries autocratizing countries.[11] Meta said that two-thirds of all influence operations it disrupted focused at least in part on domestic audiences.[12] For example, the military junta in Myanmar used Facebook as a "tool for ethnic cleansing," according to the UN,[13] and Ethiopia appeared to be on the same path.[14]

Network effects online make virality possible, giving rise to information cascades that machine learning in algorithmic content moderation and collaborative filtering systems will tend to reinforce. Coordinated propaganda campaigns take advantage of these properties of AI and the internet, for example, by targeting popular pages in hopes of reaching their audience, deploying bots to create false signals of engagement, and intimidating those with different views into silencing themselves, a hallmark of authoritarianism.

For example, in 2018, prior to the assassination of Washington Post Jamal Khashoggi by Saudi Arabia, one expert estimated that bots posted 70 to 80 percent of Arabic-language tweets containing the word "Saudi" in the previous four months.[15] Journalists were chief among the targets. Khashoggi and his colleagues had a project, the Electronic Bees Army, to frustrate Saudi influence operations by coordinating their own network of supporters that could counteract pro-government information operations.[16]

Algorithms increase the impact of the past on the present and future.[17] So the propaganda of the past is fodder for the future, recycled through unstructured datasets culled from an internet that is going to include exponentially more generative content and the digital detritus of influence operations. The strategy of undermining, drowning out, and delegitimizing real news through coordinated efforts to silence critics, flood social media, and reframe the agenda is used by digital authoritarians around the world and is recycled back into the system through natural language processing (NLP).[18]

Platforms have few economic incentives when it comes to addressing the use of their platforms for propaganda and disinformation campaigns. Influence operations are big business and have become standard operating procedure for electoral processes and public policy deliberations in all types of systems. There is no political incentive not to purchase troves of data on the electorate that will enable a campaign to micro target citizens with messages most likely to resonate with them and help them win office, pass their preferred legislation, or manufacture consent.

Fake journalist personas and think tanks are already a standard part of Russia's repertoire, carpeting online media and even respected outlets with real op-eds and analysis by faux people.[19] But now generative AI can be used to create news and other textual and visual content at a scale and scope that risks completely overrunning the systems we currently have in place to try to deal with the existing scourge of influence operations. Researchers found that the ability of recent GPT language models to generate text "at a speed, scale, and ease of use that exceeds that of troll farms and human-run disinformation campaigns" and thus represents a novel threat to the information landscape.[20]

Advances in artificial intelligence have led to the development of virtual journalists and news presenters, getting rid of the risk that a pesky journalist will go off script or interrupt a live broadcast to protest war or censorship.[21] China has already developed a bilingual AI-generated avatar news presenter for its Xinhua news agency,[22] ostensibly to reduce "news production costs" and improve efficiency.[23] The state-affiliated Abu Dhabi Media company in the Emirates has developed an Arabic speaking newscaster with the help of Chinese technology, which the company says will enable it to provide 24/7/365 news.[24] These virtual puppets are currently available in some of the more challenging digital languages, and it is a small technological leap to generate news anchors in other languages, which could pose serious issues in low digital languages and add to the arsenal of foreign influence ops. Digital authoritarians could easily overwhelm under resourced digital languages with AI-generated text that propagates their ideology while undermining or obscuring alternatives.

And it turns out that GPT-3 is already able to produce comprehensible, ideologically consistent texts in non-English, non-Latinate languages that are in line with a prompted ideological bias. It is quite good at generating extremism tests, which could be used for inculcating ideologies, cultivating propaganda, and radicalizing individuals. Researchers have shown how AI not only hallucinates,[25] but that it is possible to force GPT-3 "to integrate its innate foundation of niche knowledge with ideological bias" that can be used to create a system that responds to questions about the world with information and details that are ideologically consistent with specific worldviews.[26]

And even if generated stories don't reach their audiences, they contribute to creating skepticism and changing people's mind about critical issues, like support for a war or climate change.[27] The problem with propaganda is that it really isn't subject to factchecking because it's not about truth or falsity, it's about framing and disruption. The "ability to emulate the ideologically consistent, interactive, normalizing environment" not only makes it easier to radicalize someone,[28] but it could make it that much easier to groom the public into complacency and acquiescence.

And the risk that responses to digital authoritarianism will make it more difficult to express disagreement or opposition with the status quo poses its own threat to democracy.

Yet economic incentives are propelling the race by Google, Microsoft, and other tech behemoths to integrate ChatGPT and similar generative pre-trained transformers into search and other products is on because it is poised to become a trillion-dollar business, and there are few if any legal regulatory constraints. Microsoft has already integrated ChatGPT-3 into its Bing search engine and dozens of free and paid AI services allow rapid content creation. Once again the business models of big tech are aligned with the authoritarian potentialities of these technologies rather than their mitigation.[29]

As the power of technology and its potential for predictive manipulation grows, so does the risk that governments may choose to use it to further increase their control over public opinion and information. Generative AI technologies can be used to build personalized models for predicting people's behavior or opinions based on an individual's social media activity and other personal data, while decentralized autonomous organizations (DOAs) could be used to convey false democratic legitimacy.

For example, governments could use predictive AI and DOAs to tailor their policies, messaging, and propaganda accordingly to shape public opinion, and to burnish the appearance of democratic legitimacy through manipulated support. Furthermore, what happens when China or the UAE uses psychometric predictive AI that can generate an AI news anchor that is most likely to be perceived of as trustworthy

even as it becomes harder and harder to detect a real person from an avatar, or such efforts become irrelevant in the Metaverse?

Similarly, sophisticated algorithms can be used to analyze huge amounts of data to predict patterns of behavior or sentiment among certain groups within society. Again, this could be used by governments as a way of targeting specific individuals or groups with tailored messages to manipulate public opinion, fix elections, and alleviate pressure for reform.

Private companies also engage in predictive manipulation. Indeed, it is the business model of the surveillance economy, companies that don't use it will be less competitive, less profitable, and thus less likely to exist.[30] The economic incentives to offer faster, better, more precise predictions are myriad. So, too, are the political incentives. Preventing predictive manipulation from being used nefariously relies on little more than voluntarily goodwill.

IV. National IDs, Social Credit, and Foreclosing Civic Space

Governments that collect and analyze data from an increasing diversity of sources to monitor individuals and groups, track their activities, and build profiles on them, are creating the infrastructure upon which digital authoritarianism is built. National digital identity cards, for example are increasingly common, and allow the government to monitor and collect vast amounts of information.

The Emirates ID is a smart card that contains a microchip with biometric data, such as fingerprints and facial recognition, which are used for identification purposes. The card is required for a range of government services, including applying for residency visas, opening bank accounts, accessing healthcare. In India, a social credit system for farmers, known as the "e-NAM" (Electronic National Agriculture Market) assigns scores to farmers based on a range of factors including adherence to market rules and regulations.

The oppressive nature of surveillance is reaching new levels with the advancement of technologies such as facial recognition and predictive biometric analysis. In China and ' the UK, CCTV cameras have been outfitted with facial recognition capabilities, and law enforcement in the US and Europe can turn to companies like Clearview AI to marry their "dumb" surveillance equipment with "smart" AI systems.[31] The unregulated use of powerful biometric[32] and DNA sleuthing[33] has further enabled a constant state of surveillance that is now deeply woven into the fabric of modern life, making it almost impossible to escape the ever-watchful eye of those in power.

Expansive surveillance coupled with pervasive data collection allows digital authoritarians to develop systems that discipline and punish their citizens to keep them in line. Liu Hu, a journalist who was imprisoned after accusing a Chinese government of corruption, found that his social credit score was so damaged that he couldn't even purchase a plane ticket.[34] He was not given a chance to contest the lowering of his score, nor was he provided with any explanation for the decision. In the United Arab Emirates, authorities reportedly used the Emirates ID system to track and arrest human rights activists, monitor migrant protest leaders, and discriminate against specific types of people.[35]

Pervasive surveillance combined with the exponential advances in facial recognition and predictive biometric analysis is laying the groundwork for the destruction of anonymity in civic space. This paves the way for a future in which it is impossible to escape being constantly watched and monitored, when the idea of being able to enjoy any semblance of privacy in one's own private life is quickly becoming nothing but a distant memory.

When farmers in India staged mass protests against the government's controversial farm laws the government deployed CCTV and drones alongside social media and WhatsApp monitoring to track the movements and activities of protesters. The e-NAM system was also reportedly used to discriminate against farmers who had participated and more broadly as a tool to intimidate and silence opposition.[36]

Panoptic surveillance coupled with pervasive data collection allows digital authoritarians to develop systems that discipline and punish their citizens to keep them in line. Together with the exponential advances in facial recognition there are many cities that are laying the groundwork for the destruction of anonymity in civic space.

An example of this is social credit systems. Social credit systems use a diverse array of data to monitor and reprimand those who step out of line and punish and silence critic, anyone critical of the government, or those who belong to a particular social or political group.

The most advanced of these, China's "Social Credit System," assigns citizens a score based on a range of factors, including their financial transactions, social media activity, and behavior. Citizens with high scores are rewarded with benefits, such as access to better jobs and housing, while those with low scores may be subject to punishments, such as travel restrictions and public shaming.

The lack of transparency and due process in these systems lead to abuses of power, reinforces historic and structural deficiencies, and prevents public oversight. It also highlights how such systems can be used to further entrench authoritarianism, as those with dissenting opinions are punished and excluded from society.

The combination of a system's ability to a) track individuals' movements and monitor their online activity b) collect detailed biometric, behavioral, geographic, and financial data c) deploy facial recognition and biometric analysis in crowds and public spaces, means that privacy, anonymity, and expression can easily be curtailed, even with safeguards. Such a system is a powerful tool for surveillance and control.

It also renders the possibility of a journalist promising a source or whistleblower protection from detection obsolete, making investigative journalism more difficult and accountability less likely.

V. Trend Towards Digital Authoritarianism

Incentives to deliver public services faster and better, improve public safety, win elections, and make the most of advances in technology rarely include deterrence on the misuse of data, restrictions on how it can be used for generative AI or predictive manipulation, or transparency requirements that could help improve accountability. And some of the types of data being collected, and systems being developed, can never be used safely in a democratic system, while others threaten to overrun and drown out foundational democratic processes, from journalism to electoral process to public protest. The technological, economic, and political trends of the contemporary era trend towards digital authoritarianism since the infrastructure is being built with norms and fragile democratic institutions providing the only safeguards against their misuse.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Steve Tsou, "Council Post: The Need For Computing Power In 2020 And Beyond," Forbes (blog), January 24, 2020, https://www.forbes.com/sites/forbesbusinesscouncil/2020/01/24/the-need-for-computing-power-in-2020-and-beyond/.

2 Tsou.

3 Danielle Keats Citron, The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age (United States: W. W. Norton, 2022), https://books.google.com/books?id=MPRhEAAAQBAJ.

4 R.J. Deibert, Reset: Reclaiming the Internet for Civil Society, The CBC Massey Lectures (House of Anansi Press Incorporated, 2020), https://books.google.com/books?id=0XTqDwAAQBAJ; Ronald J. Deibert, "The Autocrat in Your IPhone," Foreign Affairs, December 12, 2022, https://www.foreignaffairs.com/world/autocrat-in-your-iphone-mercenary-spyware-ronald-deibert?check_logged_in=1."Freedom on the Net 2022: Countering an Authoritarian Overhaul of the Internet," Freedom on the Net (Freedom House, October 2022), https://freedomhouse.org/sites/default/files/2022-10/FOTN2022Digital.pdf.;

5 Steven Feldstein, "The Global Expansion of AI Surveillance," Working Paper (Carnegie Endowement for International Peace, September 2019), https://carnegieendowment.org/files/WP-Feldstein-AISurveillance_final1.pdf.

6 C. O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy (Crown, 2016), https://books.google.com/books?id=NgEwCwAAQBAJ.2016

7 Samuel Gehman et al., "RealToxicityPrompts: Evaluating Neural Toxic Degeneration in Language Models," ArXiv:2009.11462 [Cs], September 25, 2020, http://arxiv.org/abs/2009.11462.

8 "About The Pegasus Project," Forbidden Stories, accessed December 15, 2022, https://forbiddenstories.org/about-the-pegasus-project/; Boris Muñoz, "Journalism in Latin America Is Under Attack by Spyware," Wilson Center (blog), January 23, 2022, https://www.wilsoncenter.org/blog-post/journalism-latin-america-under-attack-spyware; Courtney Radsch, "AI and Disinformation: State-Aligned Information Operations and the Distortion of the Public Sphere," #SAIFE (OSCE, July 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4192038.

9 Samantha Bradshaw and Philip N Howard, "The Global Disinformation Order 2019 Global Inventory of Organised Social Media Manipulation" (Oxford Internet Institute, 2020), comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf; Craig Silverman, Ryan Mac, and Pranav Dixit, "'I Have Blood On My Hands': A Whistleblower Says Facebook Ignored Global Political Manipulation," BuzzFeed News, September 14, 2020, https://www.buzzfeednews.com/article/craigsilverman/facebook-ignore-political-manipulation-whistleblower-memo."Recapping Our 2021 Coordinated Inauthentic Behavior Enforcements," Meta (blog), January 20, 2022, https://about.fb.com/news/2022/01/december-2021-coordinated-inauthentic-behavior-report/.

10 Stephanie Kirchgaessner et al., "Revealed: The Hacking and Disinformation Team Meddling in Elections," The Guardian, February 15, 2023, sec. World news, https://www.theguardian.com/world/2023/feb/15/revealed-disinformation-team-jorge-claim-meddling-elections-tal-hanan.

11 "Democracy Report 2022: Autocratization Changing Nature?," Democracy Report (Varieties of Democracy (V-Dem) p, 2022), https://v-dem.net/media/publications/dr_2022.pdf.

12 "Recapping Our 2021 Coordinated Inauthentic Behavior Enforcements."

13 Paul Mozur, "A Genocide Incited on Facebook, With Posts From Myanmar's Military," The New York Times, October 15, 2018, sec. Technology, https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html.

14 "Hate Speech on Facebook Is Pushing Ethiopia Dangerously Close to a Genocide," accessed May 25, 2021, https://www.vice.com/en/article/xg897a/hate-speech-on-facebook-is-pushing-ethiopia-dangerously-close-to-a-genocide.

15 Elias Groll, "The Kingdom's Hackers and Bots," Foreign Policy (blog), accessed October 8, 2022, https://foreignpolicy.com/2018/10/19/the-kingdoms-hackers-and-bots-saudi-dissident-khashoggi/.

16 Katie Benner et al., "Saudis' Image Makers: A Troll Army and a Twitter Insider," The New York Times, October 20, 2018, sec. U.S., https://www.nytimes.com/2018/10/20/us/politics/saudi-image-campaign-twitter.html.

17 O'Neil, Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy.(Crown, 2016)

18 Courtney Radsch, "Weaponizing Privacy and Copyright to Silence Independent Media" (Center for International Governance Innovation (CIGI), Forthcoming); Radsch, "AI and Disinformation: State-Aligned Information Operations and the Distortion of the Public Sphere"; Carly Nyst and Nick Monaco, "State-Sponsored Trolling: How Governments Are Deploying Disinformation as Part of Broader Digital Harassment Campaigns" (Institute for the Future, 2018), https://www.iftf.org/fileadmin/user_upload/images/DigIntel/IFTF_State_sponsored_trolling_report.pdf; Courtney Radsch, "Laws, Norms and Block Bots: A Multifaceted Approach to Combatting Online Abuse," New Challenges to Freedom of Expression: Countering Online Abuse of Female Journalists (OSCE Representative on Freedom of the Media, 2016), https://doi.org/10.1163/2210-7975_HRD-4007-2016008.

19 Renée DiResta, "The Supply of Disinformation Will Soon Be Infinite," The Atlantic, September 20, 2020, https://www.theatlantic.com/ideas/archive/2020/09/future-propaganda-will-be-computer-generated/616400/.but artificial intelligence will take them to a whole new level.","container-title":"The Atlantic","language":"en","note":"section: Ideas","title":"The Supply of Disinformation Will Soon Be Infinite","URL":"https://www.theatlantic.com/ideas/archive/2020/09/future-propaganda-will-be-computer-generated/616400/","author":[{"family":"DiResta","given":"Renée"}],"accessed":{"date-parts":[["2023",3,13]]},"issued":{"date-parts":[["2020",9,20]]}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"}

20 Kris McGuffie and Alex Newhouse, "The Radicalization Risks of GPT-3 and Advanced Neural Language Models," ArXiv, September 16, 2020, 13, https://arxiv.org/abs/2009.06807. See also Noémi Bontridder and Yves Poullet, "The Role of Artificial Intelligence in Disinformation," Data & Policy 3 (2021): e32, https://doi.org/10.1017/dap.2021.20.

21 Anton Troianovski, "A Protester Storms a Live Broadcast on Russia's Most-Watched News Show, Yelling, 'Stop the War!,'" The New York Times, March 14, 2022, sec. World, https://www.nytimes.com/2022/03/14/world/europe/russian-protester-tv.html.

22 James Vincent, "China's State-Run Press Agency Has Created an 'AI Anchor' to Read the News," The Verge, November 8, 2018, sec. Tech, https://www.theverge.com/2018/11/8/18074806/ai-news-anchor-china-xinhua-digital-composite.

23 Xinhua, "World's First AI News Anchor Makes 'His' China Debut," XinhuaNet, November 8, 2018, http://www.xinhuanet.com/english/2018-11/08/c_137591813.htm.

24 WAM, "UAE Gets First Arabic Speaking AI News Anchor," Khaleej Times, April 29, 2019, sec. Tech, https://www.khaleejtimes.com/tech/uae-gets-first-arabic-speaking-ai-news-anchor.

25 Martin Anderson, "Preventing 'Hallucination' in GPT-3 and Other Complex Language Models," Unite.Ai, December 9, 2022, https://www.unite.ai/preventing-hallucination-in-gpt-3-and-other-complex-language-models/.

26 McGuffie and Newhouse, "The Radicalization Risks of GPT-3 and Advanced Neural Language Models." P. 4.

27 Thomas Marlow, Sean Miller, and J. Timmons Roberts, "Bots and Online Climate Discourses: Twitter Discourse on President Trump's Announcement of U.S. Withdrawal from the Paris Agreement," Climate Policy 21, no. 6 (July 3, 2021): 765–77, https://doi.org/10.1080/14693062.2020.1870098.Weiai Wayne Xu and Rui Wang, "Nationalizing Truth: Digital Practices and Influences of State-Affiliated Media in a Time of Global Pandemic and Geopolitical Decoupling," International Journal of Communication 16, no. 0 (January 1, 2022): 29, https://ijoc.org/index.php/ijoc/article/view/17191.

28 McGuffie and Newhouse, "The Radicalization Risks of GPT-3 and Advanced Neural Language Models."

29 Bontridder and Poullet, "The Role of Artificial Intelligence in Disinformation."

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

31 Glyn Moody, "Clearview AI Offers to Eliminate Public Anonymity and Destroy Privacy around the World for a Mere $50 Million," PIA VPN Blog (blog), February 25, 2022, https://www.privateinternetaccess.com/blog/clearview-ai-offers-to-eliminate-public-anonymity-and-destroy-privacy-around-the-world-for-a-mere-50-million/.

32 "When Bodies Become Data: Biometric Technologies and Freedom of Expression," Policy Paper (ARTICLE 19, April 2021), https://www.article19.org/wp-content/uploads/2021/05/Biometric-Report-P3-min.pdf.

33 Megan Molteni, "Cops Are Getting a New Tool For Family-Tree Sleuthing," Wired, December 16, 2020, https://www.wired.com/story/cops-are-getting-a-new-tool-for-family-tree-sleuthing/.

34 Nathan Vanderklippe, "Chinese Blacklist an Early Glimpse of Sweeping New Social-Credit Control," The Globe and Mail, January 3, 2018, https://www.theglobeandmail.com/news/world/chinese-blacklist-an-early-glimpse-of-sweeping-new-social-credit-control/article37493300/.

35 Rafeef Ziadah, "Surveillance, Race, and Social Sorting in the United Arab Emirates," Politics, September 15, 2021, 026339572110097, https://doi.org/10.1177/02633957211009719.

36 Smita Narula, "Confronting State Violence: Lessons From India's Farmer Protests," COLUMBIA HUMAN RIGHTS LAW REVIEW 54, no. 1 (2022): 82.

Protecting Digital Infrastructures, Assets, and Users From Bad Actors: An Empowering Story of Jamiiforums From Tanzania

Patricia Boshe
University of Passau*

I. The Tanzanian Context

Since independence in 1961, Tanzania has been praised as one of the peaceful countries in Africa. In 1995, the country amended the Constitution, transforming the country into a democratic state. The constitutional amendment introduced (among other aspects) a multi-party system. On the paper, this marked the end of authoritarian government and the beginning of a democratic state. However, in reality, the opposition is still relatively weak.[1] The ruling party, Chama Cha Mapinduzi (CCM) that has governed since independence continue to retain power without interruption. Even when opposition parties formed an alliance in an attempt to get rid of the one-party dominant system, they failed.[2]

The situation got worse in 2015. This is when the late President John Pombe Magufuli won the presidential election. His rule, which according to Sabatho Nyamsenda exhibited fascist symptoms, or what he choose to call authoritarian populism,[3] was characterized with intimidations, harassments, arrests and even imprisonment of political opposition and government critics; mainly to suppress freedom of expression.[4] In addition, several constitutional rights such as the right to peaceful protest and political rallying were revoked.[5] To achieve this, law enforcement used violence and physical deterrents including detention, kidnapping, and even murder.[6] Notable events include the 2018 campaign for a country-wide protest. Protests were aimed at condemning restrictions on political freedoms and the rise in human rights abuses. Planning and organization of the protests was mainly done through social media platforms. The government responded by warning citizens against participating in "the illegal and anti-government protest"[7] and dispatched police on streets across the country to deter citizens. This was also a strategy to demonstrate police capacity and readiness to resort to violence if need be.[8] In addition, those suspected of inciting and encouraging the protest were arrested.[9] Among those arrested is a group of opposition party leaders who were detained and later (in 2020) convicted for charges related to the protest.[10] As a result, the protest never took place. Freedom of expression was suppressed through a "see no evil, hear no evil" operation – a regime where the President "can do no wrong". This meant any negative publication about the President or the government would be met with negative consequences by the power that was.

The result was for the population to shift from conventional media to online media. Technology came in handy by providing the population with a "safer" space for civic participation[11], upgrading it to a primary platform for non-violence protest against

violent and non-responsive government. This was a shift that a political analyst Dan Paget believes continues to gain popularity in situations where opposition and activists are "denied access to mass media, rallies, and official posts". Meanwhile, they [opposition and activists] would double down on door-to-door politics, which could be conducted covertly. Yet their wavering members could be reluctant to venture into streets that are increasingly hostile."[12] To borrow the words of Mandira Bagwandeen, the digital space has "transform[med] the nature of contentious interactions between activists and authoritarian governments".[13]

To respond to the shift (from offline to online nonviolence activism), the government resorted to repressive laws and intrusive technological techniques to deter and maintain the control of civic activities. This fortified digital authoritarianism in Tanzania. The next section canvases the systematic ways in which laws, regulations, and technology had been employed in Tanzania to reinforce digital authoritarianism.

II. Digital Authoritarianism: Weaponizing Laws and Technology

Digital authoritarianism is essentially the use of digital technology to monitor, suppress or manipulate populations. It is a tactic mostly used by authoritarian regimes. According to Steven Feldstein, digital authoritarianism involves the use of [at least] six techniques, namely "surveillance, censorship, social manipulation and harassment, cyber-attacks, internet shutdowns, and targeted persecution against online users."[14] However, techniques to reinforce digital authoritarianism can go beyond the six, and could involve much more sophisticated means such as the use of intrusive spyware and sophisticated algorithms built on AI technologies. The ultimate goal is to gain or fortify existing controls over the population and restrict certain rights and freedoms in the digital space. According to Jason Thacker, it is a system that seeks to centralize power around political processes and individual freedoms – a system that tends to bypass constitutional accountability or oversight.[15]

Technology, like two-sided sword, provides the population with a space to exercise their civic rights free from violence but also makes it easier and swifter to centralize governing systems. In Tanzania, online surveillance (bulky and targeted), eavesdropping, real-time and geo-monitoring, and internet shutdowns are some of the techniques used to follow and identify government critics. The government then acts by intimidating, repressing, and threating them – online and with public statements. These repressive actions were used as governance measures and justified as reasonable and appropriate in maintaining national security and public good in a democratic government.[16] To give an example, the previous Deputy Minister for Transport and Communications in Tanzania, Edwin Ngonyani, insisted on monitoring social media. To him, monitoring is a necessary measure to ensure national stability and to fight against cyber-crime: "social media content has power to shape ideas and the mindset of people and if it remains uncontrolled it can lead to instability."[17] He alluded to China as the best example of countries that "managed to block such media in their country and replaced them with their homegrown sites that are safe, constructive and popular."[18]

In addition to the above methods of control, Tanzania introduced restrictive laws and regulations or amended existing laws. The objective was to create a legal framework that legalizes government control over the digital space, its content, users and creators.[19] This began with the adoption of the Cyber Crime Act of 2015 (CCA).[20] The CCA introduced offences such as publication of false, misleading, inaccurate, or damaging data, or defamatory content.[21] This law also made it an offence

for unauthorized persons to access computer data.[22] In the same year, the government adopted the Statistics Act (SA)[23] which criminalized the publishing of false official statistics and distorting facts.[24] Later, in 2018, SA was amended, introducing additional controls and criminalizing the publication of statistics without the authorization of the National Bureau of Statistics (NBS).[25] The two laws, and especially the CCA, were used to justify government's intrusion to private communications (including WhatsApp) and mobile phone tapping. Those identified as creators of such content, including anyone who forwarded them, were arrested and convicted.[26]

In 2016, the government adopted the Media Services Act of 2016 (MSA).[27] The law gave the executive wing enormous power over media and prescribed punishment for a wide range of publications. For example, section 7 (1) of the MSA gives media the right and freedom to collect, edit and publish information. However, sections 59 and 7 (2) (b) (iv) gives the Minister responsible and the government the power to prohibit and to direct the manner of reporting issues of national security and public good. In addition, through sections 36-40 and 50-53, the MSA prescribes punishment against maliciously or fraudulent fabricated information, statements threatening to national security, public safety, public order, public morals, economy of the country or those injurious to an individual reputation. In fact, a negative comment or an insult to the President is a matter of national security.[28] By 2018, within two years of coming into force of the MSA, more than 5 media houses had been fined, 7 had been temporarily shut down and 1 permanently banned.[29] To indicated how fearfully media operated in Tanzania, one newspaper: *Nipashe Jumapili*, imposed a 90-day self-ban for publishing content that displeased the President.[30] Eventually, media self-censorship became a practice.[31]

The government did not stop at content control. In 2019, it introduced an amendment[32] to the Electronic Postal and Communication Act of 2010 (EPOCA).[33] Through this amendment, the government vested more power with the Tanzania Communication Regulatory Authority (TCRA). Section 6 (of the Written Laws (Miscellaneous Amendment) Act) amended section 83 of EPOCA empowering TCRA to not only approve, but also to manage network, signal and communications equipment, including managing "their end of life processes". In 2020, the government introduced an additional regulation[34] by virtue of section 103[35] of EPOCA. The said Regulation brought restrictions on media collaborations. It prohibited content service provider hook-ups without TCRA's approval. Foreign content service providers could no longer visit or broadcast through local content provider "without being accompanied by a government official or staff from the Authority."[36] The Regulation also obliges content providers to publish their programmed schedule at least a month before it is aired.[37] Furthermore, all online content providers were to register and get licenses from TCRA. To register and get an operating license, providers must provide details of shareholders, share capital, citizenship of owners, staff qualification and training programmes, and tax clearance certificates. Failure to register and acquire a license could lead to a fine of at least 5 million shillings (approximately $2,300), jail for a minimum 12 months, or both.[38] As a result, majority of online content providers resorted to shutting down their websites.

The current regime, headed by H.E President Samia Hassan Suluhu, saw a reversal of the oppressive legal framework constructed by her predecessor, the late President Magufuli. In 2022, the government introduced the Electronic and Postal Communications (Radio and Television Broadcasting Content) (Amendment) Regulations, 2022. This regulation reversed changes introduced by the 2020 twin Regulation. In effect, a

foreign content service provider no longer needs to be approved by a government official or staff from the TCRA. Also, content providers do not need TCRA's approval to connect with other providers, and instead only need to notify the TCRA prior to the connection. In addition, section 83 of the 2022 version of EPOCA restored the original section 83 of the 2010 EPOCA. It stripped the TCRA's excessive powers. In this regard, the TCRA no longer has powers to granted by the 2019 regulation, i.e. the power to manage network, signal and communications equipment to their end-of-life processes. Moreover, the Electronic and Postal Communications (Online Content) (Amendment) Regulations, 2022[39] repealed a twin regulation that was introduced in 2020.[40] The former freed online media providers from the obligation to acquire a license; an obligation introduced by the latter to control and limit online content. As a side note, in 2018, the latter Regulation was challenged by human rights and media organzations at the High Court for restricting freedom of expression and media rights and giving TCRA wide discretionary powers in regulating content published online. The presiding judge dismissed the case on a ground that "the petitioners had failed to prove beyond doubt that the Regulations will affect their digital rights."[41]

The following section presents a story of an online forum and a man behind it who, against all odds, stood for freedoms and human rights of users of the online platform. His story is presented not only as one of a daring civic rights activist resisting authoritarianism in a seemingly democratic country in Africa, but also as an empowering story of a local platform that chose to protect people against government autocracy with very little legal backup. It's a tale of the role of civil society in promoting democratic values by resisting authoritarian practices – a heroic story of a man who not only stood between government authoritarianism and users' digital rights, but also took initiative to bring legal changes in strengthening digital rights in his country. The story is narrated in the hope that it encourages others (civil society, journalists or custodians of online content) to insist on their role in protecting digital rights.

III. Resisting Digital Authoritarianism: an Empowering Story of JamiiForums

JamiiForums (JF) (Swahili word 'Jamii' means community, therefore 'Community Forum') is a non-governmental organization based in Dar es Salaam, Tanzania. Apart from advocating for civil and digital rights, JF provides an online platform for people to freely express their opinion and access news and information. In the last decade, JF dedicated its advocacy to online safety, privacy and data protection. It was able to push the government to enact the country's first Cyber Crimes Act (the Act) in 2015. The Act is also known as "the JamiiForums Law".

In 2016, JF's founder and Executive Director, Mr. Maxence Melo Mubyazi, was arrested for his refusal to the reveal identities and personal data of whistleblowers using JF online platform to reveal government corruption and other misconduct. His arrest came after several failed trials by the government to hack, infiltrate and penetrate the platform. This failure is attributed to the fact that JF data was highly encrypted and its servers located outside Tanzania.

Like poetic justice, the government used the Cyber Crimes Act, a.k.a the Jamii Forums Law, to harass, prosecute, detain and convict Mr. Mubyazi. He was charged with obstruction of police investigation, and operating a domain not registered in Tanzania. Interestingly, three charges of obstruction of justice were open in three different courts and were being prosecuted simultaneously by different Magistrates.[42] These charges were indicted under section 22 (1) of the Cyber Crimes Act[43]. Under this section, obstructing investigation by destroying, deleting, altering, modifying

or rendering computer data useless for the purpose of investigation is a crime.[44] In two of the three cases, he was charged for failure to comply with an order to disclose data which included IP addresses of JF users.[45] According to him, complying with the request to disclose data was impractical and tainted with procedural irregularities. First, JF received multiple requests from different police departments and other institutions. This worried him as to the fate of both personal data and that of individuals identified on those requests. It also brought questions as to who the actual or intended recipient of requested information was.[46] Second, the circumstances required police to follow a procedure stipulated under section 32 (3) of the Cybercrimes Act. The section obliges police to apply to the Court for an order to disclose data. No court order was issued on the matter.

In June 2018, the first magistrate reached a decision by dismissing case No. 457. The magistrate, late Hon. Godfrey Mwambapa was then transferred to another court in upcountry Mtwara region. A second judgment was reached in April 2020 on the case No. 456. In this case, Mr. Mubyazi was convicted to a one-year prison term or a fine of 5,000,000 Tshs (approximately 2,300 US $). He paid the fine. The third judgment was reached in November 2020. Mr. Mubyazi was found guilty on charges of obstruction of police investigation, but was discharged on a warning "not to repeat the same crime within a year". The magistrate found no crime on the second charge in this case, i.e the charge of hosting the domain outside Tanzania. No law prohibited the hosting of a domain outside the country. Judgments on two of the three cases were delivered orally with no written copy to the defendants. This means, the defendants could not appeal since a written copy of the judgment must be filed with the appeal. The only written judgment was given in the case No. 456. At the conclusion of all court processes, Mr. Mubyazi had appeared before the courts for a total of 159 times.

IV. Advocating Changes for the Future of Digital Rights and Freedoms

During the time of Mr. Mubyazi's arrest and court processes, Tanzania had no comprehensive data protection law. This means, his basis for refusing to disclose data had no firm legal backup. Such legal backup is ordinarily provided by data protection laws. Regardless, he relied on article 18 of the Constitution of Tanzania which guarantees the right to privacy and JF privacy policy. To him, this brought clarity to the weakness of the country's regime in protecting individual privacy, personal data and freedom of expression. In 2021, JF pulled together human rights organizations, stakeholders from the telecommunications sector, the law society, and civil society organizations to discuss the need for a law to protect individual privacy and personal data. These organizations formed a consortium, drafted a private model personal data protection and privacy bill, and presented it to the Ministry of Information Communication and Technology (ICT). The Ministry accepted the proposal for the law. It then drafted and tabled its draft bill to the Parliament. This was in November 2022. The law has been published on the parliament website, and the Ministry is now working on developing its regulations. This was an activity involving the consortium of stakeholders formed by Mr. Mubyazi. He believes this law is one of the cornerstones that supports freedom of expression.

V. Democratic Reversals

A combination of restrictive legislations, threats, and bans on individual freedoms, placed the state of Tanzania democracy in question. The situation also demonstrates that laws can be (mis)used to suppress individual freedoms and rights. During the time when Tanzania implemented suppressive laws, the population lost their freedom to speak, to assemble and exercise civic rights, to political expression, of

media, and to private communication and private life. Instead of protecting individual rights, these laws "legalized" oppression and control of the population by the government.

The situation in Tanzania shows how fast a democratic state can change into an autocratic state. It also shows how laws and technology can facilitate and empower authoritarianism. The story of JF and Mr. Mubyazi illustrates the role of civil organizations in holding governments accountable, promoting democratic values, as well as protecting individual rights and freedoms. Most importantly, it reminds society to be proactive. In responding to political situations, we should seek out both technological and administrative approaches to secure rights and freedoms of the population served. JF provided a platform for people to speak freely about the government. JF was aware that government whistleblowers used the platform to uncover government malpractices. This made JF a target of government surveillance and hacking activities. To protect its users, JF implemented privacy and data protection standards. They used highly sophisticated encryption, which blocked the government from accessing the identities of platform users. Locating its servers outside Tanzania prevented the government from illegally accessing and harvesting personal data of these users. On one hand, JF users were protected against autocratic government acts, while on the other hand JF provided the Tanzania population with the only platform to practice their civic rights, speak freely and even question government conduct.

VI. Recommendations

The description of what happened in Tanzania provides us with a few points to take home. Supporting civil and human rights organizations / activists in their roles as guardians and mouthpieces of the rights and freedoms of the people means:

1. Implementing robust digital security – it is necessary to conduct an impact and situational assessment and put your houses in order – and have an appropriate (based on situation and the level of technological advancement) security measure such as:

    a. encryption to secure privacy and protect data of your digital platform users,

    b. having a privacy and consumer protection policy published and consumer protection policy published – let users know the extent of protection you offer so that they may decide on the manner and extent of their interaction on your online platform.

2. Proactivity in legal and policy changes. Monitor policy and legal changes, identify loopholes that may be used to diminish users' rights and freedoms or your ability as a platform owner / content provider to uphold and protect the rights and freedoms of your users. Involve other stakeholders and provoke conversation on loopholes. When possible, engage the government in legal reforms or policy changes.

3. Create a Plan B. In countries where internet shutdowns are a "norm", civil society, online platform providers or content creators should devise an emergency offline or secured network to continue holding the government accountable. Such alternatives should not wait until internet shutdown occurs. It should be created, rehearsed and proofed (functionality) prior to internet shutdowns.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 https://freedomhouse.org/search?key=tanzania

2 Dan Paget "Tanzania: The Authoritarian Landslide" Journal of Democracy, vol. 32, no. 2, 2021, pp. 61–76 at 66; See also Dan Paget "Reinterpreting Authoritarian Populisms: The Elitist Plebeian Vision of State" Political Studies, Online publication https://journals.sagepub.com/doi/10.1177/00323217231154098

3 Sabatho Nyamsenda "Bulldozing like a Fascist? Authoritarian Populism and Rural Activism in Tanzania" International Institute of Social Studies (ISS) in The Hague, Netherlands Conference Paper No.78, ERPI 2018 International Conference Authoritarian Populism and the Rural World 17-18 March 2018.

4 Ibid.

5 Isaac Mugabi "Tanzanian President Ends Ban on Opposition Rallies" Deutsche Welle 01/06/2023; https://www.dw.com/en/tanzania-president-samia-suluhu-hassan-scraps-opposition-meeting-ban-imposed-by-magufuli/a-64303922

6 "Tanzania: Criticism of President Leads to Imprisonment -Two Tanzanian opposition political leaders are sentenced to prison amidst growing concern for Tanzania's commitment to democracy and human rights" Media House Press Release 27/02/2018.

7 "Tanzanian President Warns Against Illegal Anti-Government Protests" African News, 10/03/2018.

8 Abdur Rahman Alfa Shaban "Tanzania Police Ready to Cripple Defiant Anti-Govt Protesters" Reuters African News 22/03/2018.

9 "Tanzania: Authorities Seek to Muffle Protests: Tanzanian Authorities Pre-emptively Arrested Citizens using Social Media to Promote the Protest on April 26" Media House Press Release 24/04/2018.

10 "Tanzania Opposition Leaders Found Guilty for 'Illegal' Protest" the Monitor 11/03/2020; See also Press releases Office of the High Commissioner for Human Rights "UN Rights Chief Disturbed by Harassment of Opposition following Tanzania Elections" 10/11/2020 https://www.ohchr.org/en/press-releases/2020/11/un-rights-chief-disturbed-harassment-opposition-following-tanzania-elections

11 Dan Paget "Tanzania: The Authoritarian Landslide" p.179.

12 Dan Paget "Tanzania: The Authoritarian Landslide" p. 179.

13 Mandira Bagwandeen "Don't blame China for the Rise of Digital Authoritarianism in Africa" Blog post 09/09/2021 https://www.fpri.org/article/2021/09/dont-blame-china-for-the-rise-of-digital-authoritarianism-in-africa/

14 Steven Feldstein, When it Comes to Digital Authoritarianism, China is a Challenge — But Not the Only Challenge, February 12, 2020 Commentary

15 Jason Thacker What is digital authoritarianism? Human Dignity, Technology/October 28, 2020 https://jasonthacker.com/2020/10/28/what-is-digital-authoritarianism/

16 Emma Santana Fano "Digital Authoritarianism in Sub-Saharan Africa. Internet Control Techniques and Censorship: A Qualitative-comparative Analysis" Master's Thesis, 2020.

17 Asterius Banzi "Tanzania Seeks Chinese Help in Social Media" The East African Newspaper, 01/08/2017.

18 Ibid.

19 A similar approach is used in Russia. Russia also has legislations that requires content producers to register with the government and holds them liable for site content if they fail to block secured networks maneuver such as the use of vpns.

20 Act No. 14 of 2015.

21 Section 16 reads: Any person who publishes information or data presented in a picture, text, symbol or any other form in a computer system knowing that such information or data is false, deceptive, misleading or inaccurate, and with intent to defame, threaten, abuse, insult, or otherwise deceive or mislead the public or councelling commission of an offence, commits an offence, and shall on conviction be liable to a fine of not less than five million shillings or to imprisonment for a term of not less than three years or to both.

22 Section 7

23 Act No. 9 of 2015.

24 Section 37 (5) of Act No. 9 of 2015.

25 Acts No. 8 of 2018.

26 Section 20 of the CCA prohibits the sending or transmitting of unsolicited messages. The penalty upon conviction is a fine of not less than three million shillings (approximately 1300 $) or three times the value of undue advantage received, whichever is greater or to imprisonment for a term of not less than one year or to both. See section 20 (2) CCA.

27 Act No. 12 of 2016.

28 Charlotte Cross "Cybercrime and the Policing of Politics in Tanzania" in Maggie Dwyer / Thoman Molony (Eds.), Social Media and Politics in Africa: Democracy, Censorship and Security (London: ZED 2019) pp. 195-213 @ p. 202.

29 Mwananchi Newspaper 15/01/2018; The Guardian Newspaper 03/11/2017.

30 Ibid.

31 According to a report published by Freedom House, "Unsurprisingly, the crackdown [on rights and freedoms] has had a chilling effect on independent news sources. Widespread self-censorship is apparent in the conspicuous lack of local coverage of scores of bodies washing up on the beaches of Dar es Salaam, the reported disappearance of over a thousand children in the Kibiti area, and the extrajudicial killings of young men by watu wasiojulikana ("unknown persons"). The disappearance of a journalist in November 2017 gave the media another reason to avoid reporting on sensitive topics." "A Cautionary Tale for Tanzanian Democracy Activists" Media House 24/04/2018 https://freedomhouse.org/article/ethiopia-cautionary-tale-tanzanian-democracy-activists

32 This was through the Written Laws (Miscellaneous Amendment) Act, No. 5 of 2019.

33 Act No. 3 of 2010.

34 The Electronic and Postal Communications (Radio and Television Broadcasting Content) (Amendment) Regulations, 2020.

35 This section allows, on the one hand, the Minister to make media content related regulations, and on the other, TCRA to make media content related rules.

36 Regulation 4 of the Electronic and Postal Communications (Radio and Television Broadcasting Content) (Amendment) Regulations, 2020. This Regulation amended Regulation 37 of the Electronic and Postal Communications (Radio and Television Broadcasting Content) Regulations, 2018. It inserted a sub regulation to that effect.

37 Regulation 5 of the Electronic and Postal Communications (Radio and Television Broadcasting Content) (Amendment) Regulations, 2020 read together with regulation 39 of the Electronic and Postal Communications (Radio and Television Broadcasting Content) Regulations, 2018

38 Tanzania's communications regulator had given bloggers, as well as owners of other online forums such as YouTube TV channels, until May 5 to heed tough new internet content rules through state registration and a license fee of up to $900. See Fumbuka Ng'wanakilala "Tanzania Sets Two-Week Deadline for Bloggers amid Internet Crackdown" Reuters 24/04/2018.

39 Government Notice No. 136 published on 18/3/2022.

40 The Electronic and Postal Communications (Online Content) Regulations, 2020.

41 High Court of Mtwara, Miscellaneous Case No. 25 of 2018.

42 Republic v. JamiiForums, case no. 456 of 2016; Republic v. JamiiForums, case no. 457 of 2016; and Republic v. JamiiForums, case no. 458 of 2016.

43 Case No. 458 on managing a domain not registered in Tanzania was brought under section 79 (c) of EPOCA

44 The crime is punishable by fine of not less than three million shillings (approximantely 1300 US $) or to imprisonment for a term not less than one year or both.

45 These are cases No. 456 and 457.

46 He wrote this on JF platform for his readers. See https://www.jamiiforums.com/threads/hukumu-dhidi-yangu-shukrani-kwa-watanzania-maxence-melo.1712418/

Part III
Global Digital Finance and Democratic In/Exclusion

In this section:

- **Digital Finance and the Specter of Digital Authoritarianism**
- **Undoing Democratic Social Citizenship? The Digitalization of G2p Payments and the Making of Private Digital Authoritarianism**

Berggruen Institute | Canada | UCLA Institute for Technology, Law & Policy | USC Dornsife Center on Science, Technology, and Public Life

Digital Finance and the Specter of Digital Authoritarianism

Saule T. Omarova
Cornell University*

I. Digitization of Money as a Political Phenomenon

The term "digital authoritarianism" refers generally to "the use of digital information technology by authoritarian regimes to surveil, repress, and manipulate domestic and foreign populations."[1] It embodies the "dark" side of technological progress that enables new forms of oppressive government action on an unprecedented scale. To date, much of the academic and policy literature on digital authoritarianism has focused on two technologically savvy regimes actively using digital tools for domestic and foreign policy purposes: China and Russia.[2] As more authoritarian governments around the world adopt similar strategies of technologically-driven political control and manipulation, it increases the need for democratic societies both to understand these dynamics and to protect themselves from these new harms.

Digital authoritarianism, however, is a complex phenomenon that cannot be reduced to overtly political government actions. Private entities that develop, own, or control certain types of digital technology and market infrastructures are increasingly poised to become both the enablers of authoritarian politics and potential sources of new, more invisible forms of oppression. Dominant market actors can weaken democracy by wielding their economic power in a way that increases their private gains at the expense of the long-term public interest, political integrity, and participatory values.[3] Unchecked concentrated private power is an existential threat to democratic political order.

These dynamics are increasingly obvious in modern finance. In developed democracies, the financial system is the principal arena for generation and allocation of money and credit - the universal production input and a critical infrastructural resource in any decentralized exchange-based economy. Financial markets are the key distributional and governance mechanisms that support and shape modern democratic societies. Behind its seemingly apolitical technical facade, finance is an inherently political phenomenon: those who control where the money flows effectively control whose voice counts. A democratic society cannot survive without a financial system that can be trusted to create and allocate money and credit in a democratic and fair fashion.

Technology is a crucial factor in this respect. It can be used either to enhance the trust in and democratic potential of our financial system, or to destroy it by enabling few dominant actors to usurp and abuse financial power. In the last decade, the digital "revolution" in financial services has forced this tension into the spotlight.[4] Yet, there is still little understanding among policymakers and public interest advocates of the potentially game-changing *political implications* of this process. Discussing the evolving digital finance in the context of digital authoritarianism helps to highlight these implications.

Digital currencies provide a useful starting point in this inquiry. Digitization of money offers well-known transactional benefits: faster, cheaper, direct peer-to-peer payments that can be automated. Through technological design and cryptography, digital money offers its users an ability to avoid reliance on traditional banks and other centralized authorities, promising both greater payment security and freedom from state corruption and control.[5] Since the launch of Bitcoin, which embodied this fundamentally libertarian paradigm, however, the evolution of cryptocurrencies has shown the limits of decentralization in today's complex, scale-driven financial system. The most ubiquitous digital currencies that perform key money functions in crypto-markets – including so-called "stablecoins" – are issued and managed by private entities that run centralized transactional platforms for these currencies. Moreover, Big Tech conglomerates and large Wall Street banks, able to leverage their existing customer bases and large balance sheets, are steadily expanding into the markets for digital assets. Finally, government actors are also increasingly involved in digital finance, indirectly through regulation and directly through adoption or issuance of digital currencies.[6] In short, behind the rhetoric of decentralization and democratization, there is a struggle for control of the rapidly transforming financial and economic markets – and the polities deeply dependent on them.

To explore these dynamics, I focus on two closely linked issues: (1) direct use of digital currencies by authoritarian governments; and (2) the rise of private digital money as a challenge to democratic governance. The goal of this paper is to highlight the principal ways in which technology widely touted for "democratizing" finance can, in fact, undermine the existing democratic governments and facilitate a global shift toward a more hierarchical and exploitative economic and political order. While limited in scope and detail, this analysis helps to expose the emerging specter of digital money as *anti-democratic* medium. It also reveals the fundamentally political trade-offs behind many of today's seemingly technical policy choices. I argue, accordingly, that policymakers in developed economies should proactively and deliberately address both of these problems in an integrated manner, as a multi-faceted *structural threat* to democratic government.

II. Digital Currency as a Tool of Authoritarian Governments

While government actors can adopt or promote domestic use of any existing cryptocurrency (e.g., Bitcoin or Tether), issuing its own digital money is a more direct way to harness the power of technology for political purposes. In modern financial systems, central banks can issue digital versions of sovereign money: "central bank digital currency" (CBDC). In contrast to both traditional bank deposit-money and private cryptocurrencies, CBDC is a digitally represented direct monetary liability of the central bank (or an equivalent monetary authority). Currently, such digital sovereign liabilities exist only in the form of special "reserve" account balances held at the central bank by banks and, in some jurisdictions, certain other licensed financial institutions. Depending on the design, CBDC can expand access to this safest-of-the-safe, fully public electronic money beyond banks and other privileged financial institutions. Doing so can make the monetary system genuinely inclusive and free of private profit-gaming incentives.

A growing number of central banks around the world are experimenting with CBDC.[7] At present, however, none of the central banks in developed democracies (except, perhaps, for Sweden) are close, or have committed, to issuing CBDC in the near future.[8] The most frequently stated concern in this respect is the potential impact of centralizing digital money issuance in the government's

hands on individual freedom and privacy. Another, less often vocalized but extremely powerful obstacle to CBDC adoption is so-called "disintermediation" of private banks. The banking industry's exclusive ability to fund itself through publicly-insured retail deposits is a tremendous economic advantage that the industry seeks to amplify through digital technology. As the safest and most liquid digital asset in a monetary system, CBDC can displace, or even replace entirely, private banks' deposit-money. In countries with developed private financial sectors, in which large-scale financial investment and trading markets are functionally connected to and dependent on bank credit, this would cause significant structural shifts. The financial industry thus strongly opposes any CBDC plans that, regardless of potential public benefits, supplant or impinge on private money-creation.

By slowing down the process of digitizing sovereign money, however, domestic interest-group politics can leave developed democratic countries' vulnerable in the international arena. Unlike pluralist democracies, authoritarian regimes are not hesitant to use CBDC as the means of political domination, both internally and externally. Because CBDC can move on different payment rails than the traditional sovereign currencies, it can be weaponized, among other things, to bypass existing international banking channels in order to evade economic sanctions or to reconfigure global alliances. Not surprisingly, both China and Russia have publicly committed to digitizing their currencies.[9]

China is the undeniable leader in the CBDC race, the only major economy with a "live" CBDC in the public testing phase. The People's Bank of China (PBOC) has been actively working on CBDC since 2014 and launched its first multi-province "digital yuan" (e-CNY) pilot in early 2020.[10] The Chinese government has partnered with private Chinese banks and technology companies to offer mobile apps, ATMs, and other supporting services for e-CNY. It also offered direct money incentives for retail users, to encourage e-CNY adoption on a scale sufficient to get the system ready for use at the at the 2022 Beijing Winter Olympics.[11] By the end of August of 2022, transactions using e-CNY surpassed $13.9 billion (or 100 billion yuan). The spending involved 360 million transactions in pilot areas in 15 provinces and municipalities, with more than 5.6 million merchants accepting digital yuan payments.[12] In 2022 alone, participating regions have reportedly offered 30 rounds of e-CNY subsidies – including the $4.5 million in free e-CNY drop in Shanghai – designed to stimulate consumption and serve other macroeconomic purposes.[13] During the Lunar New Year period, the government distributed nearly $30 million dollars' worth of e-CNY to boost its use for holiday spending and gifting.[14] As reported in early 2023, China is planning to expand its digital yuan pilot to most of its 1.4 billion population.[15]

From the start, the Chinese government has approached the e-CNY issuance as a long-term strategic project serving multiple political goals. As designed, the e-CNY system gives the PBOC sole direct access to users' financial and transactional data and ability to monitor individual payments in real time. This new form of mass surveillance qualitatively expands the familiar toolkit of digital authoritarianism and amplifies its capacity to collect information far beyond its own borders.[16] For example, at the Beijing Winter Olympics, the e-CNY was reportedly used to make more than $315,000 of payments a day, which raised concerns about the possibility of enabling Chinese government to monitor foreign citizens' financial transactions.[17] Centralized state control of money and payments can also be used as an instrument of coercion, as the government can punish political dissenters by freezing or expropriating their funds and effectively

cutting them off from the financial system. It can also manipulate people's opinions and behavior by "rewarding" political loyalty via direct credit of their e-CNY wallets or by subsidizing intentionally designed "activities." The malleability and programmability of digital money makes it a particularly effective instrument of power.

The e-CNY is also closely tied to China's geopolitical ambitions. Despite China's leading role in international trade, its currency currently constitutes a tiny share of international payments settled primarily in USD. Issuing its own digital currency will enable the Chinese government to create an alternative global payments system it can effectively control. Coupled with the Chinese government's long-standing strategy of investing in the infrastructure and acquiring real assets in foreign countries, the globalization of the digital yuan will solidify China's economic dominance and expand its geographic sphere of influence. It will also insulate China and its geopolitical allies from economic sanctions and other means of political pressure exerted by the international community.[18] This technologically-enabled concentration of financial, infrastructural, and political power in China's hands will fracture the global system and potentially subject many sovereign countries' populations to mass surveillance and oppressive tactics deployed by the Chinese government against its own citizens.

At this relatively early stage, it is difficult to identify the full range of dangerous implications of using CBDC as a tool of digital authoritarianism, In part, the picture may be hazy because the Chinese government has been working on the e-CNY rollout in partnerships, both globally with other governments and domestically with private companies. In 2022, the PBOC partnered with central banks of Thailand, Hong Kong, and United Arab Emirates to test a multi-CBDC cross-border payments platform, m-Bridge, where the e-CNY was the most actively used settlement currency.[19] Domestically, PBOC issues and distributes its e-CNY not directly but through several state-approved banks. The digital yuan is accepted for payments on China's major social media and e-commerce platforms, including WeChat and Alibaba that offer their own competing payment apps.[20] Private companies' participation may help to obscure the nature and extent of the government's control over the emerging ecosystem, thus legitimizing the entire project as an innovative economic experiment. It is critical to remember, however, that the public-private dynamics in digital finance are bound to reflect the underlying balance of public and private powers in the relevant political system.

III. Private Stablecoins as a Challenge to Democratic Sovereignty

In developed economies with democratic governments, digitization of money has been primarily a private market phenomenon, whereby multiple privately-issued digital currencies co-exist, sometimes uneasily, with sovereign money. The recent growth of stablecoins poses an especially serious challenge to the long-term stability and monetary sovereignty of democratic governments. Stablecoins are digital currencies claiming to keep "stable" value pegged to the value of the USD or some other state currency. Typically, the issuer of a stablecoin – including USDC, Tether, and Binance USD –maintains the peg to traditional money by setting up a "reserve" to hold USD or other safe assets backing it. A stablecoin thus "borrows" its stability from the sovereign money and functions as its *tokenized derivative*, or a privately-controlled digital representation.[21]

In this sense, stablecoins are both a direct competitor to, and a direct outgrowth of, sovereign currency. They facilitate trading, lending, and investing in a wide variety of crypto-assets, particularly in the so-called

decentralized finance (DeFi) universe, and serve as "onramps" connecting crypto-markets to the traditional financial system.[22] This critical infrastructural function gives stablecoin issuers – private crypto-exchanges, banks, and technology companies – potentially enormous market power. As the issuers of the widely accepted "means of exchange" and "store of value" within the crypto ecosystem, these private market actors can replicate the functions of traditional central banks, but without the accompanying legal obligation to act in the public interest. In effect, stablecoins enable what may be called *synthetic privatization* of the fundamental public function and a critical public resource – sovereign money and credit – with no democratic accountability or express political commitment to provide public goods.

The political risks this business model creates are especially visible in the case of Big Tech stablecoins. In June 2019, Facebook (now Meta) launched its Libra project (later renamed Diem): a global stablecoin to be issued by a Swiss-based consortium of large corporations led by Facebook. From the start, Libra was promoted as a service for the billions of people around the globe locked out of the traditional financial system.[23] The original plan was to have the Libra Association issue a global cryptocurrency, backed by a basket of sovereign currencies (the "Libra Reserve"), and manage a cross-border payments network built on top of Facebook's vast social-media platform. Facebook was to run the digital wallet built into the Libra ecosystem, and provide other potentially lucrative services tied to it.[24] The unprecedented scale and structural design of this project, which would effectively put Facebook at the center of global money and payments flows, generated strong political and regulatory backlash. Despite the newly renamed Meta's efforts to scale down and rebrand it, the project was ultimately wound down.[25]

The Libra/Diem saga is nevertheless highly instructive. It revealed how private digital currencies and payments systems created and controlled by globally dominant techno-financial conglomerates can directly threaten the stability, autonomy, and resiliency of the world's leading democracies. If successful, the Libra/Diem stablecoin would have made Meta a "shadow" Federal Reserve, a source of globally ubiquitous digital currency, potentially more powerful than the actual Fed.[26] It would have given Meta and its corporate partners direct access to the financial and transactional data generated by the users. Meta would have been able to monitor in real time daily activities of billions of users, manipulate their preferences and shape their behavior, and otherwise commercialize their personal data in deeply invasive ways. It would have been able to condition individuals' access to, or price of, Libra/Diem either on their willingness to purchase other goods or services offered by Meta or on some form of "social scoring" maintained by it.[27] From there, it is not hard to imagine Meta using its newly-minted power over digital finance for political reasons, in effect replacing old interest-group politics with the new tactics of digital authoritarianism.

This is not such a far-fetched scenario, particularly given the highly personality-driven culture of the tech and crypto industries. Successful technology firms – including publicly-traded Apple, Microsoft, Amazon, and Meta – tend to be closely associated with their charismatic founders. Many tech firms also have corporate governance structures that explicitly concentrate control in the hands of the few insiders.[28] These private authoritarian tendencies under the guise of techno-meritocracy are even more extreme in crypto-finance, where they provide the fertile ground for the rise of potentially autocratic "visionaries" with grand political ambitions.

The tech-driven ability to control digital money and finance networks, therefore, offers not only an unprecedented economic advantage but also a new set of previously non-existent *political levers*.

From this perspective, the ongoing expansion of Big Tech platform companies into digital money and payments – including the recent announcement of Twitter's intentions to offer payments services[29] – raises potentially far more troubling and complex issues than is commonly acknowledged. The same is true of financial institutions, such as JPMorgan that currently issues its own JPM Coin and runs a permissioned blockchain platform for trading of tokenized financial instruments.[30] It is critical to see these private conglomerates' push into digital money not simply as a technologically innovative business strategy, but as a politically salient project of redefining the core public-private balance in finance.[31]

To date, the prevailing response to this challenge has been a call for regulation. What the appropriate regulatory regime should seek to accomplish, and by what means, however, remains unclear. In the U.S., the vast majority of policy proposals seek to make private stablecoins actually "stable" and "safe" as the means of payments, either by limiting their issuance to federally-insured banks or by mandating the composition of reserves backing them.[32] While consumer protection and avoiding failure are important policy goals, this approach ignores or downplays the perilous structural implications of legitimizing – and publicly subsidizing – private stablecoins. Instead of defending the state's monetary sovereignty, this seemingly pragmatic approach risks further erosion and ultimate loss of democratic control of public money and finance.

But there is a bigger problem with treating regulation as the only possible response. As shown in my prior work, our existing *technocratic* paradigm of financial regulation is inherently ill-suited to deal with the unique challenges of digital finance.[33] That raises the question about other, more direct and effective, options we may need to consider.

IV. Reclaiming Control: CBDC as a Democratic Project

This brief overview of China's CBDC project and private stablecoins brings into a sharp relief the frequently overlooked political link between these phenomena, both of which implicate (directly or indirectly) concerns with the rise of digital authoritarianism. Policymakers in developed democracies need to respond to this double-threat decisively and in an integrated manner.

As a practical matter, the United States and other democratic countries need to elevate CBDC issuance to the top of their current policy priorities. Our policymakers, industry leaders, and public interest advocates should start treating CBDC (digital USD, Euro, British pound, etc.) not as simply "one of many potential means of digital payments" – the common view that frames the CBDC debate in narrowly technocratic terms – but as the means of ensuring the sovereign public's continuing control over its economic and political affairs, globally and domestically.

Globally, digitizing the US dollar and developing a viable network of interoperable CBDCs is vital to avoiding a potentially irreversible shift in the international balance of power in favor of China and its allies. Democratic governments cannot outsource this global "monetary countervailing" role to private stablecoin issuers, relying on their presumably superior innovation capacities or "natural" free-market preferences.

Wall Street and Big Tech conglomerates are driven by private profits and have strong incentives to transact in the digital yuan or the digital ruble, if doing so is sufficiently profitable. Private firms cannot supplant state actors in the international arena, and the advent of CBDC extends this logic to global payments.

Operationalizing this logic requires much closer cooperation and coordination among the central banks in key jurisdictions (United States, Canada, European Union, United Kingdom, Japan, Australia, etc.), which need to ensure seamless flow of sovereign digital currencies across jurisdictional borders and markets. In addition to ramping up the ongoing technical work on cross-border CBDC interoperability, it is important to actively consider more ambitious and potentially durable political solutions. These may include establishing regional digital currency unions or even the creation of a supranational CBDC – the digital-era version of Keynes's "bancor" and the International Clearing Union.[34]

It is also essential to establish new modes of coordination and cooperation among each country's own central banks, financial regulators, and foreign policy and national security agencies. At the national level, the ongoing work on CBDC design and adoption needs to be an integral part of a broader government strategy, so that foreign policy objectives and concerns can continuously inform, contextualize, and shape the technical or regulatory decisions. Expanding CBDC designers' mindset beyond the immediate financial market dynamics is an important step toward "getting it right."

Domestically, CBDC issuance appears increasingly necessary in order to preserve and strengthen the role of *public money* in the digital era. The proliferation of private digital assets, including stablecoins, does not obviate the fundamental need for public money but threatens to push it down into the dark "basement" of the financial system, where its only function would be to backstop private digital money. If allowed to happen, the gradual disappearance of public money from economic transactions will severely weaken the vital connection between the state and the citizenry, effectively incapacitating the former and leaving the latter without guaranteed access to fully safe monetary instruments.

Central bankers and policymakers need to recognize CBDC as the best available defense against this internal threat. They should broaden their view of how CBDC can be used, and what public benefits it can generate, outside of the narrow sphere of digital payments. Rather than vilifying state-issued digital currency, policymakers (in collaboration with scientists, legal experts, and public interest advocates) need to prioritize development of technological, legal, and institutional safeguards of CBDC users' financial privacy and autonomy.[35]

Finally, policymakers need to expand the range of CBDC design options beyond what is currently on the table, or what would be least disruptive from the private financial industry's perspective. Instead of allowing the industry lead the process, policymakers need to formulate a coherent agenda for using CBDC as a tool of sustainable and equitable growth, democracy, and prosperity.[36] CBDC has tremendous potential to redefine the central bank balance sheet as the ultimate public platform for the creation and allocation of money and credit to productive economic enterprise. Whether or not we realize this potential may determine the future of our democracy in the era of digitization.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Alina Polyakova and Chris Meserole, "Exporting Digital Authoritarianism: The Russian and Chinese Models," Brookings Institute Policy Brief, 2019, 2, https://www.brookings.edu/research/exporting-digital-authoritarianism/.

2 Ibid.

3 "Facebook – Cambridge Analytica data scandal," Wikipedia, visited February 17, 2023, https://en.wikipedia.org/wiki/Facebook%E2%80%93Cambridge_Analytica_data_scandal.

4 For in-depth analysis, see Saule T. Omarova, "New Tech v. New Deal: Fintech as a Systemic Phenomenon," Yale Journal on Regulation 36 (2019): 735-93.

5 Archie Chaudhury, "Reflecting on Satoshi Nakamoto's Manifesto," Bitcoin Magazine, October 31, 2022, https://bitcoinmagazine.com/culture/reflecting-on-satoshi-white-paper.

6 "Central Bank Digital Currency Tracker," Atlantic Council, accessed February 28, 2023, https://www.atlanticcouncil.org/cbdc-tracker/.

7 Anneke Kosse and Ilaria Mattei, "Gaining momentum – Results of the 2021 BIS survey on central bank digital currencies," BIS Papers No. 125, May 6, 2022, https://www.bis.org/publ/bppdf/bispap125.htm.

8 In 2019, the Bahamas became the first country to issue its Sand Dollar. Swedish Riksbank has been piloting its eKrona since 2017, due to the declining use of cash. Katharina Buchholz, "Where Central Banks Have Issued Digital Currencies," Forbes, July 15, 2022, https://www.forbes.com/sites/katharinabuchholz/2022/07/15/where-central-banks-have-issued-digital-currencies-infographic/?sh=b9094c5513e1.

9 Helen Partz, "Russia to roll out CBDC pilot with real consumers in April," Cointelegraph, February 17, 2023, https://cointelegraph.com/news/russia-to-roll-out-cbdc-pilot-with-real-consumers-in-april.

10 Nathaniel Popper & Cao Li, "China Charges Ahead with a National Digital Currency," New York Times, March 1, 2021, https://www.nytimes.com/2021/03/01/technology/china-national-digital-currency.html. The e-CNY is officially known as the Digital Currency Electronic Payment (DCEP).

11 Ibid.

12 "China's digital currency passes 100 bln yuan in spending – PBOC," Reuters, October 13, 2022, https://www.reuters.com/markets/currencies/chinas-digital-currency-passes-100-bln-yuan-spending-pboc-2022-10-13/.

13 Ibid.

14 Jesse Coghlan, "China doles out million in digital yuan in bid to boost adoption," Cointelegraph, February 6, 2022, https://cointelegraph.com/news/china-doles-out-millions-in-digital-yuan-in-bid-to-boost-adoption-report.

15 Marc Jones, "New look CBDCs and cryptomarket to emerge from turmoil," Reuters, February 6, 2023, https://www.reuters.com/technology/new-look-cbdcs-cryptomarket-emerge-turmoil-top-bis-official-says-2023-02-06/.

16 James Kynge and Sun Yu, "Virtual control: the agenda behind China's new digital currency," Financial Times, February 17, 2021, https://www.ft.com/content/7511809e-827e-4526-81ad-ae83f405f623.

17 Marc Jones, "Over $315,000 in digital yuan used daily at Olympics, PBOC official says," Reuters, Feb. 16, 2022, https://www.reuters.com/technology/around-300-mln-digital-yuan-used-every-day-olympics-pboc-official-says-2022-02-15/.

18 Georgina Lee and Samuel Shen, "China's digital yuan stands out in cross-border pilot in a show of global ambition," Reuters, October 27, 2022, https://www.reuters.com/markets/currencies/chinas-digital-yuan-stands-out-cross-border-pilot-show-global-ambition-2022-10-27/.

19 Ibid.

20 Arendse Huld, "China Launches Digital Yuan App – All You Need to Know," China Briefing, September 22, 2022, https://www.china-briefing.com/news/china-launches-digital-yuan-app-what-you-need-to-know/.

21 Bank for International Settlements, "The Future Monetary System," BIS Annual Economic Report, June 21, 2022: 78, https://www.bis.org/publ/arpdf/ar2022e3.pdf.

22 Ibid.: 80-81.

23 Julia Boorstin, "Facebook launches a new cryptocurrency called Libra," CNBC, June 18, 2019, https://www.cnbc.com/2019/06/17/facebook-announces-libra-digital-currency-calibra-digital-wallet.html.

24 Ibid.

25 Olga Kharif, "Meta-Backed Diem Association Confirms Asset Sale to Silvergate," Bloomberg, January 31, 2022, https://www.bloomberg.com/news/articles/2022-01-31/meta-backed-diem-association-confirms-asset-sale-to-silvergate?leadSource=uverify%20wall.

26 Saule Omarova and Graham Steele, "There's A Lot We Still Don't Know About Libra," New York Times, November 4, 2019, https://www.nytimes.com/2019/11/04/opinion/facebook-libra-cryptocurrency.html.

27 Ibid.

28 Michele Corgatelli, "Tech Companies and the Attractiveness of Dual Class Shares," Fordham Journal of Corporate & Financial Law Blog, Aug.27,2021, https://news.law.fordham.edu/jcfl/2021/08/27/tech-companies-and-the-attractiveness-of-dual-class-shares/.

29 Hannah Murphy, "Elon Musk pushes forward with Twitter payments vision," Financial Times, January 30, 2023, https://www.ft.com/content/9d84d534-b2dd-4cff-85d1-aee137b26a45.

30 "Onyx by J.P.Morgan," J.P.Morgan, accessed February 28, 2023, https://www.jpmorgan.com/onyx/coin-system.htm

31 Omarova, "New Tech v. New Deal," 790-92.

32 David L. Portilla, Danjie Fang, "An Overview of the Stablecoin Policy Debate," ABA Business Law Today, October 10, 2022, https://www.americanbar.org/groups/business_law/publications/blt/2022/10/stablecoin-policy-debate/.

33 Saule T. Omarova, "Technology v. Technocracy: Fintech as a Regulatory Challenge," Journal of Financial Regulation 6 (2020): 75-124.

34 James M. Boughton, "Why White, Not Keynes? Inventing the Postwar International Monetary System," IMF WP/0252, March 2002, https://www.imf.org/external/pubs/ft/wp/2002/wp0252.pdf; Yanis Varoufakis, "Imagining a new Keynesian Bretton Woods," Project Syndicate, May 6, 2016, https://www.weforum.org/agenda/2016/05/yanis-varoufakis-imagining-a-new-keynesian-bretton-woods.

35 Raul Carrillo, "Seeing Through Money: Democracy, Data Governance, and the Digital Dollar," Georgia Law Review, forthcoming 2023, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4354085.

36 Saule T. Omarova, "The People's Ledger: How to Democratize Money and Finance the Economy," Vanderbilt Law Review 74 (2021): 1231-1300.

Undoing Democratic Social Citizenship? The Digitalization of G2p Payments and the Making of Private Digital Authoritarianism

Nick Bernards
University of Warwick*

I. Rise of Social Transfer Programs

One of the most significant, if sometimes overlooked, transformations in the nature of democratic citizenship over the last four decades has been the rise of social transfer programmes across much of the global South. While often obscured by the rising precarity and erosion of some important labour protections amidst neoliberal reforms, the past 40 years have in fact witnessed an *expansion* of flagship social security programmes in the global South, albeit often through a variety of different forms of non-contributory social transfers rather than conventional contributory social security.[1] Notable examples include Brazil's *Bolsa Familia* programme of conditional cash transfers and South Africa's interlocking grid of Child, Old-age, and Disability transfers. Kevin Harris and Ben Scully, invoking Polanyi, call this a 'hidden counter-movement', noting that the expansion of these programmes has been driven by social and political struggles on the part of precarious workers themselves, often facilitated by political openings created by democratization.[2] Rina Agarwala similarly points to the growing extent to which 'informal' workers in India direct grievances towards the state and in terms of citizenship, given that they often do not have clearly identifiable employers against whom to make claims.[3] Tania Li more generally points to the increasingly central role played by a politically articulated 'sense of entitlement' in struggles over poverty and inequality.[4] In short, social transfers are quietly perhaps the most important fruits of the so-called 'third wave' of democratization, and undoubtedly part of the political bargain that has allowed the neoliberal package of privatizations, labour market flexibility, financial deregulation and the like to coexist with multiparty democracy in much of the world.

We can point, in short, to an emerging form of 'democratic social citizenship' distinct from classical understandings of social citizenship. T.H. Marshall's landmark intervention on 'citizenship and social class' pointed to the emergence of a model of 'social citizenship' in which political membership was increasingly being associated with more extensive rights to social and economic security.[5] In Marshall's articulation, these entitlements remained intimately linked to particular modes of work -- secure, unionized, and well-paying jobs backed by generous pension and social security -- held up as 'standard' in Western Europe and North America in the mid-twentieth century. Marcel Paret rightly notes that in this sense Marshall's sociology of citizenship was reflective

of a geographically and historically exceptional context.[6] Even in their heyday, these articulations of secure livelihoods and citizenship themselves were always deeply inflected with racial, colonial, and gendered hierarchies.[7] The emerging forms of social citizenship described above do not resemble the mid-twentieth century forms charted by Marshall. They are without doubt limited in important ways.[8] But they are democratic forms of social citizenship nonetheless, especially insofar as they link political membership and democratic participation to a suite of social and economic entitlements.

Of late, the matrices of social transfers sitting at the core of new forms of democratic social citizenship are also increasingly targets for digitalization. Digitizing social transfers is held out by advocates as a means of making payments faster, cheaper, more efficient, and reducing fraud and waste, and (no less importantly) of promoting other ancillary goals like expanding the use of formal financial channels. My argument in what follows is that this project of digitalization is being carried out on terms that threaten to undermine emerging modes of democratic social citizenship, and inaugurate new privatized digital authoritarianisms. The digitalization of social transfers, or 'G2P' payments in the new jargon, represents an important potential shift in the character of state-citizen relations in peripheral democracies. It is one that threatens to undermine key elements of actually-existing democratic social citizenship, and to foreclose the expansion of the latter.

II. Clashing Visions of Social Protection

As Christopher Webb aptly observes, the pandemic has brought to the fore clashing visions for the future of social protection.[9] Early on, governments rushing out emergency aid measures for people affected by lockdowns raised hopes of a transformation of social policy. Civil society groups and organized labour in global North and South sought to press for expanded forms of social protection funded through progressive taxation on the back of these measures. Meanwhile, in Webb's words, we can also see the growing prominence of a 'coalition of financial institutions, technology companies and development institutions who see this as an opportunity to further the inclusion of the poor into financial markets'. The World Bank and major donors seized on the pandemic to push for the closer integration of mobile and digital payment systems, and digital identity verification systems, with public systems for social protection.

Shortly before the outbreak of the pandemic, the World Bank had launched the 'G2Px' initiative, together with the Gates Foundation, aiming to encourage wider adoption of digital payment and digital identity systems in administering state transfer payments in the global South. The outbreak of COVID-19 presented a singular opportunity to press this project forward. Digital payments were framed in the early days of the pandemic as means of delivering emergency social assistance while minimising physical contact. They were also more generally presented as a way of facilitating more rapid distribution of emergency aid. World Bank officials argued that 'countries with advanced G2P payment ecosystems are able to push transfers out with lightning speed'.[10] By contrast, 'In countries where investments in payment infrastructure and DFS have not yet been made and where regulations have not been modernized, scaling up G2P and continued access to financial services will be more difficult'.

Yet G2Px is clearly aiming for more thoroughgoing change to social protection systems than simple faster delivery of state transfers. In at least two ways. The digitalisation of G2P payments is very explicitly seen by the Bank and others as a way of expanding the use of digital financial services in general. The Bank's officials have made little secret that the wider purpose of digitising emergency G2P payments is to ramp up adoption of digital financial tools more broadly: 'the crisis may represent an opportunity to fast track changes already in the works in areas such as interoperability and mobile money adoption and DFS in general'.[11] Equally, as I'll show further in the next section, G2Px and like projects imagine social transfers as a hyper-targeted, residual 'safety net' to be used primarily to manage emergencies rather than, say, a mechanism for redistribution.

The digitization of G2P payments threatens to create new private digital authoritarianisms, in at least two ways. First, digitalization efforts often also encode particular approaches to social protection, based on the rapid disbursal of highly targeted emergency assistance, often with the primary aim of increasing participation in the formal financial system, over universal forms of social provision. Second, digitalization efforts have often led to the privatization by stealth of significant portions of social protection systems, leading to entrenched private sector monopolies that are difficult to displace, and the erosion of state capacity.

III. Removing Social Transfers from Democratic Challenge

The focus on targeting, the blurring of 'financial inclusion' objectives with social protection, and emphasis on using social transfers to expand 'access' to financial services ultimately threaten to entrench a particular set of assumptions about the aims and objectives of social transfers. Deliberately or not, projects of digitalization in their current form threaten to entrench and depoliticize, and hence remove from democratic challenge, social transfer systems designed around some quite specific and limited aims.

Much of the work on G2Px emphasizes the ways that digital identification, in conjunction with machine-learning and other tools, might enhance public authorities' ability to target the delivery of social assistance.[12] In one flagship report for the programme, the World Bank notes that one of the chief benefits of adopting digital identification procedures in G2P payments is that 'the ability to cross-check various databases using the unique identifier has made it possible to increase the precision of targeting'.[13] The digitisation of social assistance, then, is tied directly to a particular vision of social protection based on the rapid, targeted delivery of emergency cash assistance to those most in need, coupled with a constant vigilance for fraud, waste, and duplication.

The blurring of social protection with financial inclusion should also give us pause. As noted above, the wider adoption of digital financial services is increasingly spoken about in G2P debates as an end in itself. This is a problem. Not only do the benefits of financial inclusion for poverty reduction remain ambiguous and contested,[14] but prioritizing financial inclusion entails to a considerable degree entrenching a set of objectives quite different from those of democratic social citizenship in social transfer systems.

It's also worth noting the new kinds of costs such programmes might impose on targeted populations. If mobile and digital payment systems are indeed faster and cheaper for banks and governments, this is in no small part because they offload some processing and transaction costs onto users' devices.

At bare minimum, using a mobile or digital payment system to access social assistance requires access to a mobile phone and airtime. Marco Haenssgen, writing in the context of mobile health applications, describes the compulsion for marginalized people to buy and maintain mobile devices in order to maintain access to health services as establishing a kind of 'tyranny of technology adoption'.[15] Much the same arguably applies to social transfers if people have no choice to but to pay for phones in order to access key services or entitlements. The Bank's discussions of G2P payments often hint at an awareness of some of the possible exclusionary impacts of greater reliance on digital finance to deliver social assistance. But responses thus far have often centered on behavioural science-themed efforts to alleviate 'mistrust' of digital services.[16] Tying social transfers to 'financial inclusion' aims also at least opens the door for new kinds of predatory and exploitative practices -- as shown further in the South African case below.

Deliberately or otherwise, then, the digitalization of social transfers -- at least in the form pursued through G2Px -- threatens to displace emergent forms of democratic social citizenship into limited, inflexible, even exploitative privatized forms. It threatens, in short, to create new privatized digital authoritarianisms, combining disciplinary surveillance and hyper-targeted emergency aid, and redirecting the aims of social transfers away from meeting the entitlements of democratic social citizenship and towards expanding participation in the formal private sector. Digitalization also threatens to create privatized payments infrastructures which are difficult to displace.

IV. Case Study: Net1 in South Africa

The history of South African payments firm Net1 is instructive.[17] Net1 is a South African fintech company, and a pioneer of biometric cash transfer services. By 2012, through a subsidiary named Cash Paymaster Services (CPS), it was responsible for the administration of virtually all of South Africa's expansive system of social grants. Notably, Net1 was heavily supported by the World Bank through the Bank's private-lending arm, the International Finance Corporation, and was frequently cited as a model of digitally-enabled social provision that encouraged wider financial access.[18]

Net1 and CPS' role in administering the social grants quickly became increasingly controversial, both because of actual or alleged irregularities in the tendering process, and because it was linked to facilitating exploitative lending practices. Net1 leveraged its near monopoly control over the distribution of social payments through CPS in order to aggressively market loans to transfer recipients. Net1 used its position managing the flow of social transfers into recipients accounts to deduct loan payments from transfer payments directly.[19] As Erin Torkelson puts it, 'Funded by the state, Net1 turned social grantees into a lucrative and risk-free market'.[20]

While the South African government came under increasing pressure to drop Net1's contracts to deliver social grants, particularly after a Constitutional Court ruling in 2014 that those contracts had been tendered illegally, it took until 2018 to replace Net1. Breckenridge stresses that Net1 was able to 'lockin' the state's dependence on its services in order to deliver social grants: 'For millions of the most vulnerable people—especially those living in the poorest regions of the country—only Net1 can deliver the grants'.[21]

Digitisation of social protection, in short, runs the risk of entrenching the role of private finance and technology firms in ways that entrench exploitative practices and undercut the capacity of the state, making them very difficult to undo.

Problems of this kind linked to the digitalisation of welfare state systems are not confined to developing countries, either. As Rosie Collington has recently documented in detail with respect to Denmark, the digitalisation of social protection schemes has gone hand in hand with piecemeal processes of privatisation and retrenchment. Digitisation has been accompanied by the transfer of critical administrative infrastructures to private actors. Paradoxically, while digitalisation has typically been justified in terms of improved 'efficiency' and speed, 'public sector capacity, and hence the ability to achieve public goals, has been undermined'.[22]

V. Digitalizing Social Protection?

Digitalizing some aspects of social protection administration might under some circumstances be beneficial. Whether this is the case, however, is profoundly contingent on questions of ownership, capacity, and of basic aims. Philip Alston, the Special Rapporteur to the UN Human Rights Council on Extreme Poverty and Human rights, has argued that digitalization is often used as a 'trojan horse' for welfare retrenchment.[23] He suggests, usefully, that 'Instead of obsessing about fraud, cost savings, sanctions, and market-driven definitions of efficiency, the starting point should be on how existing or even expanded welfare budgets could be transformed through technology to ensure a higher standard of living for the vulnerable and disadvantaged'.

In its present form, the digitization of social transfers carries a very real threat of creating new privatized digital forms of authoritarianism -- undermining or circumscribing emergent forms of democratic social citizenship, and entrenching private monopolies over the delivery of significant state functions.

It ultimately matters a good deal who owns and provides the plumbing for social payments infrastructures. A payments infrastructure which is controlled by a handful of monopolistic firms, designed around principles of surveillance and targeting, and which shunts key costs of operation onto users is likely to do more harm than good. As the case of Net1 in South Africa shows especially clearly, ceding control over key infrastructures to private companies makes them exceedingly difficult to change or displace. Digital social protection systems must, as a starting point, be publicly owned in order to be democratically controlled.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 A. Barrientos, M. Niño-Zarazúa and M. Maitrot, Social Assistance in Developing Countries Database, Version 5.0 (Manchester: Chronic Poverty Research Centre, 2010).

2 K. Harris and B. Scully, "A Hidden Counter-Movement? Precarity, Politics and Social Protection Before and Beyond the Neoliberal Era", Theory and Society 44, no. 5 (2015) pp. 415-44.

3 R. Agarwala, Informal labour, formal politics and dignified discontent in India (Cambridge: Cambridge University Press, 2013).

4 T.M. Li, "After development: Surplus populations and the politics of entitlement", Development and Change 48, no. 6 (2017), pp. 1247-1261.

5 T.H. Marshall, Citizenship and Social Class (Cambridge: Cambridge University Press, 1950).

6 M. Paret, "Citizenship and work in global capitalism: from domination to aspiration", Sociology Compass 12, no. 8 (2018) pp. e12602.

7 R. Shilliam, Race and the Undeserving Poor (Newcastle: Agenda, 2018); L. Vosko, Managing the Margins: Gender, Citizenship, and the International Regulation of Precarious Work (New York: Oxford University Press, 2010).

8 See, for instance, A. Saad-Filho 'Social policy for neoliberalism: The Bolsa Familia programme in Brazil', Development and Change 46, no.6 (2016), pp. 1227-1252.

9 C. Webb, "Financial inclusion and the future of social protection policy", Developing Economics, available: https://developingeconomics.org/2021/08/30/financial-inclusion-and-the-future-of-social-protection-policy/, accessed 28 February 2023.

10 M. Rutkowski, A. Garcia Mora, G.L. Bull, B. Guermazi, and C. Grown, "Responding to crisis with digital payments for social protection: Short term measures with long-term benefits", World Bank Blogs, available: https://blogs.worldbank.org/voices/responding-crisis-digital-payments-social-protection-short-term-measures-long-term-benefits, accessed 28 February 2023.

11 Ibid.

12 See e.g. E. Aiken, G. Bedoya, A. Coville, J.E. Blumenstock, 'Targeting development aid with machine learning and mobile phone data: Evidence from an anti-poverty intervention in Afghanistan', COMPASS '20: Proceedings of the 3rd ACM SIGCAS Conference on Computing and Sustainable Societies June 2020, pp. 310–311 -- co-authored by a pair of officials at the World Bank and two researchers from UC Berkeley.

13 World Bank, Scaling up social assistance payments as part of the COVID-19 pandemic response (Washington: World Bank Group, 2020).

14 See, inter alia, P. Mader, 'Contesting Financial Inclusion', Development and Change 49, no. 2 (2018), pp. 461-483; N. Bernards, A Critical History of Poverty Finance: Colonial Roots and Neoliberal Failures (London: Pluto Press, 2022).

15 M. Haenssgen, 'The struggle for digital inclusion: Phones, healthcare and marginalisation in rural India', World Development 104 (2018), p. 371.

16 World Bank and Ideas 42, 'Behavioural Science for Inclusive and Impactful Digital Payments', Product Design Case Study (2021), available: http://documents.worldbank.org/curated/en/213761626156031874/Behavioral-Science-for-Inclusive-and-Impactful-Digital-Payments, accessed 28 February 2023.

17 See K. Breckenridge, 'The global ambitions of the biometric anti-bank: Net1, lockin and the technologies of African financialization', International Review of Applied Economics 33, no. 1 (2019), pp. 93-118.

18 E. Torkelson, 'The World Bank's role in SA's social grants system', GroundUp (2017), available: https://www.groundup.org.za/article/world-banks-role-sas-social-grants-payment-system/; accessed 28 February 2023.

19 This practice is discussed at length in E. Torkelson, 'Sophia's choice: Debt, social welfare, and racial finance capitalism', Environment and Planning D: Society and Space 39, no. 1 (2021), pp 67-84.

20 Ibid, p. 68.

21 Breckenridge, 'global ambitions', op cit., p. 94.

22 R. Collington, 'Disrupting the welfare state? Digitalisation and the retrenchment of public sector capacity', New Political Economy 27 (2): 312-328.

Part IV
New Tech: Despotic Blockchains and Exploitative AI

In this section:

- **Technocratic Despotism in the Network State**
- **Platform Authority and Data Quality: Who Decides What Counts in Data Production for Artificial Intelligence?**

Berggruen Institute | Canada | UCLA Institute for Technology, Law & Policy | USC Dornsife Center on Science, Technology, and Public Life

Technocratic Despotism in the Network State

Morshed Mannan
European University Institute*

I. Digital-First States

In *The Network State* (TNS), a book by the influential technology entrepreneur and investor Balaji Srinivasan, the author envisions the creation of digital-first states with blockchain technology serving as its "backbone".[1] Network States, Srinivasan's argument goes, would emerge from "a highly aligned online community with a capacity for collective action that crowdfunds territory around the world and eventually gains diplomatic recognition from pre-existing states".[2] There are at least two major reasons why policy makers and scholars concerned about the rise of digital authoritarianism should pay attention to the concept of The Network State.

First, there are key differences between TNS and earlier libertarian exit projects. TNS has attracted sizable interest online, particularly in the short period since the eponymous book was published.[3] The view that a state-building project must first involve cultivating a community of people aligned around a moral innovation before the acquisition of land already marks a departure from earlier libertarian projects like Roatán Próspera in Honduras[4] or Satoshi Island in Vanuatu.[5] The latter group of projects focused on land acquisition first, and community building second. They sought to acquire land and enact favorable legislation to build private "charter cities" and crypto-economies to attract a global well-heeled community to a "crypto-utopia".[6] Many of these projects fizzled out or never gained momentum in the first place. By focusing on community-building around one cause or moral premise—what Srinivasan calls "one commandment"[7]—TNS is able to tap into anxieties about global crises, and the decline of communities and trust across a wider range of people than crypto-enthusiasts.[8] Conversely, TNS is also distinct from The Silk Road—a dark net platform with a libertarian community and social norms—that has also been compared to a (proto-) state as TNS does not merely advocate withdrawal from the state, but actively seeks to acquire territories within existing states.[9] The appeal of the ideas espoused by TNS can be seen in initiatives like the Praxis Society, which plans to initially develop a community of like-minded individuals before acquiring territory on the Mediterranean coast.[10] There is still financial interest in libertarian exit projects,[11] so the growth of network states can be anticipated in the coming years.

Second, beyond its nascent popular appeal, academics, particularly those subscribing to agorist counter-economics[12] or belonging to the school of Austrian economics,[13] have written positively about the benefits of being able to choose between competing governance and regulatory systems through territorial (i.e., leaving a territory) and non-territorial exit (i.e., switching to another system without leaving a territory). The opportunities afforded by the use of encryption technologies associated with blockchains to escape

detection by nation-states is seen as being particularly promising. "Cryptosecession" can realize the benefits of people being able to switch political jurisdictions as easily as changing their IP addresses with a VPN, and gradually removes "items from the bundle of government goods and services, and reassign[s] them to diverse, increasingly non-territorial public and private enterprises".[14] Whether successful or not, in the view of authors like MacDonald, cryptosecession projects can curb the alleged "fiscal exploitation" of states and improve legacy systems.[15]

This paper, after providing a brief overview of the concept of TNS, presents a critical analysis of how this nascent political project conceives state building and demonstrates the potential it has for creating a form of technocratic despotism. In short, it institutes a governance system that, at first blush, provides near-limitless choice but in reality enables the unchecked exercise of power by technical experts, including in abusive ways. In lieu of exit-based governance and opt-in consent, this paper concludes by stressing the need for voice-based, multistakeholder governance for digital communities. I now turn to a brief overview of The Network State.

II. The Network State

TNS builds on the premise that nation-states are neither good nor are they reformable; with it being easier to start a new state than engage in the politics necessary to change existing ones.[16] These new states provide blank canvases for economic and governance systems that incumbent states are not equipped to provide and technological innovations that they would not permit. This is why, in his view, network states are needed.

The book defines the concept as follows:

> "*A network state is a* ***social network*** *with a* ***moral innovation****, a sense of* ***national consciousness****, a* ***recognized founder****, a capacity for* ***collective action****, an* ***in-person level of civility****, an integrated* ***cryptocurrency****, a* ***consensual government*** *limited by a* ***social smart contract****, an archipelago of* ***crowdfunded physical territories****, a* ***virtual capital****, and an on-chain census that proves a* ***large enough population, income, and real-estate footprint*** *to attain a measure of* ***diplomatic recognition.****"*[17]
> [emphasis added]

A network state is the result of a succession of overlapping steps and milestones. The network Srinivasan has in mind begins with a group—a "startup society"— that is socially connected, digitally networked, and geographically dispersed. The members of this network *opt-in* to join but entry is not completely free of any conditions and he countenances the possibility of expulsion for "bad behavior".[18] The network is organized around a moral purpose rather than profit. This is because, like libertarian founders such as Michael Oliver before him, he sees "missionary societies" as being more likely to survive than "mercenary" societies.[19]

Yet, a single moral innovation is not enough, as it is only through feelings of sharing other similar values and being able to act on these values through group agency that a startup society can gradually develop into what Srinivasan calls a "network union".[20] This is why he stresses the need for a national consciousness and the capacity to act collectively. The reader should not be under any illusions that this collective action will be bottom up and non-hierarchical, as "[a] state, like a company, needs a leader...it's important to have a recognized founder, one that people actually listen to and choose to follow by joining the

community".[21] Channeling Carl Schmitt, the author contends that only recognized leaders can resolve contentious disputes and more democratic arrangements cannot. Community members consent to this governance arrangement by the very act of signing a "social smart contract" and agree to give this leader "*limited* privileges over the user's digital life".[22] This leader will ensure that the purpose of the community will be preserved and subsequent steps are taken to transform the network union into a network state. How will the leader be held accountable? Through the opportunity to *exit*. While he claims that network states will have a "decentralized administration", his vision of its governance appears to be distinctly techno-authoritarian as it comprises a founder/chief executive and their engineers, who write prescriptive and proscriptive laws in code, and have them enforced through cryptography.[23] To pre-empt accusations of opening the door to abusive leadership, he argues that an oppressive network would not be able to gain or retain citizens.

However, the very issue with digital communities, even if passionate about a certain subject, is that they are fluid and have a constant churn of members, enabled by this very ability to exit. Trust needs to be built within the community through, for instance, in-person meetings. While allowing for exit, to proceed in the development of a network state there needs to be a means of objectively and convincingly authenticating the membership and assets of a community at any given time. An integrated cryptocurrency and its underlying blockchain become important in this regard as it can serve as the tamper-resistant, constantly updated record-of-truth of the digital state and its virtual capital, for uses ranging from acting as a means of payment and exchange to the administration of records, registrations, and asset ownership to identity verification and community censuses. Cryptocurrency is thus essential for network states as it won't be possible to have true sovereignty without a sovereign digital currency.

As a core community of people coalesces, and accumulates assets, it is possible to transition from being a network union to a "network archipelago",[24] by crowdfunding the acquisition of properties in different parts of the globe. However, the idea is not just limited to acquiring properties ranging from apartments to islands, it is also to physically lock out those who do not belong or those who have lost their privileges. Srinivasan uses the example of a door that can only be unlocked with their ENS name, i.e., their blockchain-native username associated with their Ethereum public address.

Finally, if a network archipelago has gathered enough clout through its acquisition of aligned people, capital, and real estate it is possible and essential for it to negotiate diplomatic recognition from other states. For Srinivasan, diplomatic recognition is what makes network states "real"[25] and it is the only possible means to prevent invasion from other states at will.

His ultimate example of a network state is a US Food and Drug Administration (FDA)-free zone where people can try to develop life-enhancing biomedicine. It would need diplomatic recognition so that its authority to be exempt from the laws of the FDA are not compromised by other states. Bitcoin (i.e., the peer-to-peer currency and payment system) and blockchain technologies are not only essential to this project, they are seen as making its success inevitable, as they cannot be shut down and can continue to operate in areas that are outside of the control of states. In the following section, I present three main criticisms of the concept of the network state.

III. Critique of the Network State

Firstly, recreating states in the image of technology startups leads to unjust legal and governance systems.[26] "Cryptostatecraft" is not an uncommon aspiration among techno-libertarians[27] as others have also had fantasies of nation-states splintering into "gov-corps" ruled by CEOs.[28] For its proponents, cryptostatecraft is desirable as developers ostensibly want to build and politicians want to exert power over others.[29] However, cryptostatecraft ignores the nature of power that would exist in these communities, the forms of (venture capital) finance that would exert influence on its prerogatives, and the sinister implications this startup-centric view holds for governance and law enforcement. In this vision, it would be an opaque 'Network' that would be the Leviathan that ensures humans in the network state act in prosocial ways, rather than God or the State.[30] While the Network Leviathan is presented as a fusion of cryptography and community (with a murky system of governance), it is the function of Bitcoin that is most pronounced: through its very existence as a technology that cannot be fully constrained by any one state it becomes a "government of governments", providing people with alternatives to governments that cause inflation and seize private property.[31] By design, communities are supposed to enjoy the freedoms of "Pax Bitcoinica" and, unlike in real liberal democratic states, have no say in its rule or capacity to question its merits. Implicitly, this regime silences dissent and difference of values.[32] The absence of voice mechanisms is especially pernicious as network states will continue to have the same "division of labor" needed for capitalism to function,[33] but without the countervailing forces that are needed to prevent exploitation. As a result, the concept of citizenship, which in practice has already become contingent on the value of marketable skills in some countries,[34] will be further commodified and decoupled from political rights.

In network states, "[l]aw enforcement", will be able to "flip digital switches as necessary to maintain or restore domestic order, just like the sysadmins of today's tech companies".[35] Due process, the rights of representation, and other hard-fought protections —conceded even by Locke in describing spontaneous, voluntary associations[36]—-would seemingly be eliminated upon entry into this jurisdiction of a network state. The experience of charter city projects like Roatán Próspera, where the Honduran Constitution and local laws were changed to grant these projects significant regulatory autonomy from the state, shows that such a political jurisdiction is possible.[37] Through the techno-solutionist hype of network states, even more physical space will be converted into what Kitchin and Dodge call 'Code/Space', in which "software and the spatiality of everyday life become...produced through one another".[38] In other words, physical space will not be usable as intended without properly functioning code. While this might create new conveniences for a few—like the 'internet-enabled' modular homes being built on Satoshi island—it will also create new zones of exclusion for people who do not have the power or privilege to dream of network states.

Second, opt-in consent is an inadequate basis for legitimate governance. According to TNS, the fact that people would freely and voluntarily consent to join a network state is to be proof that its governance is legitimate and that it can exercise power over people. This understanding of legitimacy is predicated on a contractualist understanding of legitimacy which can arise simply through tacit consent.[39] There are three reasons for questioning this claim. Firstly, freedom of contract and association are not without limits, and even in the contracts

that Srinivasan analogizes network state contracts with—employment contracts, among others—there are contracts that are void because their substance is illegal (e.g., slavery) or violates principles of non-discrimination.[40] Secondly, potential citizens may not understand the terms of their "social smart contract" and thereby enter into arrangements they did not expect.[41] Imagine, for instance, a network state formed to support refugees: does a refugee who cannot go anywhere else give actual quality consent to, for example, having their irises scanned each time they want to share a medical record with a health professional? It is possible that they have no other choice but to consent, even when they do not understand the full implications of such data sharing.[42] Tacit oral consent or even the use of private keys to sign a transaction would appear to fall short of actual quality consent in such a situation. Thirdly, examples such as these show how remaining within a particular territory, should not in itself be read as tacit agreement to obey a particular social contract as they would lead to unjust outcomes: most people have no choice but to remain as there is nowhere to go.[43] The consensual government of TNS, given the nugatory role of deliberation, would have no space for giving refugees a voice in setting the terms of the social smart contract or amending it. Instead, when challenged, as with earlier voluntary platforms with ideological goals like The Silk Road, the use of opt-in consent takes on authoritarian undertones: "Whether you like it or not, I am the captain of this ship...[Ross Ulbricht, the founder of The Silk Road wrote to a disgruntled user]...You are here voluntarily, and if you don't like the rules of the game, or you don't trust your captain, you can get off the boat".[44] In such situations and many others like it, the choice to opt-in may be an illusion and instead be a form of coercive control.

Third, exit-based governance has significant limitations. To explain the benefits of exit, Srinivasan presents the example of *runxue*, a movement of disaffected Chinese citizens who study ways to leave China following the imposition of COVID lockdown measures. We can see that exit has a protective function (i.e., by enabling Chinese citizens to act in self-defense), an expressive function (i.e., by demonstrating their discontent), and an individual autonomy-enhancing function (i.e., by allowing them to self-determine their own lifepath).[45] Friedman adds that there are other advantages of exit-based governance, such as exit being less epistemically demanding on individuals compared to voice as they don't need to speculate on the implications of a governance problem on others or consider the best potential solution.[46] Those exiting only need to rely on local and relatively reliable knowledge that they experience themselves.[47]

However, even advocates for exit-based governance concede that for it to be feasible, there needs to be some comparable space to escape to, and even more importantly, it requires an extensive form of redistribution that would make exit affordable—something "far more ambitious than a universal basic income".[48] As Craib powerfully puts it: "Preferential exit under radically unequal conditions, is not benign withdrawal in the pursuit of autonomy and self-government but a continuation of class warfare by other means".[49] As redistribution is not countenanced by TNS, and there are significant costs in starting a startup society, in all likelihood exit will only be between startup societies created by rich founders.[50]

Despite exit being considered by many techo-libertarians as the only fundamental right humans should have,[51] TNS offers little guidance on the terms and conditions of exit. As citizens of network states have

no say in setting the rules of governance, or the rules to change rules,[52] restrictive conditions and penalties could be set on exit without its outright prohibition.[53] A 'price' of exit might be, for instance, a non-disclosure agreement forbidding discussion about the reasons for exiting or the activities of the network state. The particular circumstances of each network state can also affect the feasibility of exiting from certain activities. Some activities do not lend themselves to exit as their results can't be iteratively corrected elsewhere—like risky biomedical procedures that can be potentially lethal.[54] Other activities can't be exited easily because of the time-commitment and costs involved, making it infeasible to start again elsewhere.

The opportunity to escape online into an open metaverse thus offers little comfort as it is an option open to relatively few. Yet, should a network state fail in achieving its purpose, it is those founders and members who have powerful passports that will be able to readily abandon the network state, leaving those less fortunate on the hook with nowhere to go.

IV. Alternative Imaginaries

Network states, like the metaverse before it, are 'hyperstitional', in the sense that their very existence as ideas help generate the activity and resources that bring it into being.[55] However, concepts such as TNS risk bringing into being a form of technocratic despotism because of the three reasons mentioned in the previous section. There are at least five domains, operating at different scales, in which the battle to institute voice over exit-based governance needs to be fought. These are: the geopolitical domain (i.e., between nation-states and between nation-states and networks), the protocol domain (i.e., the technical layer of networks), the nation-state domain, the civil society domain, and the business domain. Active and inclusive multi-stakeholder governance, drawing on a wide variety of innovations by multi-stakeholder networks, cooperatives and commons,[56] could be drawn on to inform the voice-based mechanisms in each domain. Various technologies could be deployed to enable multi-stakeholder governance, including blockchains, but this alternative would not elevate any technology to the status of globe-spanning Leviathan.

The book, *Networked Governance of Freedom and Tyranny*, offers valuable lessons on multi-stakeholder governance in networks in the context of state-building in newly independent countries like Timor-Leste. The authors show how polycentric network governance among many actors was crucial to securing Timor-Leste's freedom, but to uphold civic republicanism, these networks needed constant renewal and a more extensive separation of powers beyond the traditional three organs of state. Feminist networks and other indigenous institutions had important functions in securing these checks-and-balance.[57] Such approaches are not—as charged by Friedman—epistemically demanding as they rely on local knowledge, yet the use of voice rather than exit allows for this knowledge to be shared and better decisions to be made. Research that explores alternative imaginaries to network states such as communations/coordi-nations is underway, and should receive closer attention from those interested in solutions to the rise of digital authoritarianism.[58]

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Balaji Srinivasan, The Network State: How to Start a New Country (1729, 2022), 223.

2 Srinivasan, 9.

3 Balaji [@balajis], 'How do we get the first network state with diplomatic recognition? We'll need a pipeline of hundreds, perhaps thousands of startup societies.', Tweet, Twitter, 20 July 2022, https://twitter.com/balajis/status/1549780617659039745.

4 Raymond Craib, Adventure Capitalism: A History of Libertarian Exit, from the Era of Decolonization to the Digital Age (Oakland, CA: PM Press, 2022)

5 "About Satoshi Island", Satoshi Island, accessed February 25, 2023, https://www.satoshi-island.com/#about.

6 Jillian Crandall, 'Blockchains and the "Chains of Empire": Contextualizing Blockchain, Cryptocurrency, and Neoliberalism in Puerto Rico', Design and Culture 11, no. 3 (2 September 2019): 280, https://doi.org/10.1080/17547075.2019.1673989.

7 Srinivasan, The Network State, 136.

8 Douglas Rushkoff, Survival of the Richest: Escape Fantasies of the Tech Billionaires (New York: W.W. Norton, 2022); Raghuram Rajan, The Third Pillar: How Markets and the State Leave the Community Behind (New York, NY: Penguin, 2020); Robert Putnam, Bowling Alone: The Collapse and Revival of American Community (New York: Simon & Schuster, 2000).

9 Rita Zajácz, 'Silk Road: The market beyond the reach of the state', The Information Society 33, no. 1 (2017): 23, http://dx.doi.org/10.1080/01972243.2016.1248612.

10 "The Plan", Praxis Society, accessed February 25, 2023, https://www.praxissociety.com/plan.

11 Harrison Smith and Roger Burrows, 'Software, Sovereignty and the Post-Neoliberal Politics of Exit', Theory, Culture & Society 38, no. 6 (1 November 2021): 143, https://doi.org/10.1177/0263276421999439.

12 Nikos Sotirakopoulos, 'Cryptomarkets as a libertarian counter-conduct of resistance,' European Journal of Social Theory 21, no. 2 (2018): 189, https://doi.org/10.1177/1368431017718534

13 Trent J. MacDonald, The Political Economy of Non-Territorial Exit: Cryptosecession (Kindle ed., Cheltenham: Edward Elgar Publishing, 2019); Jeffrey Friedman, Power Without Knowledge: A Critique of Technocracy (Oxford: OUP, 2019); Aviezer Tucker and Gian Piero de Bellis, Panarchy: Political Theories of Non-Territorial States (Abingdon: Routledge, 2015).

14 MacDonald, The Political Economy of Non-Territorial Exit, 113-114. Also see, Darcy W.E. Allen, Chis Berg, and Mikayla Novak, 'Blockchain: an entangled political economy approach', Journal of Public Finance and Public Choice 33, no. 2 (1 October 2018): 105, https://doi.org/10.1332/251569118X15282111163993.

15 MacDonald, 347; Craib, Adventure Capitalism, 8.

16 Srinivasan, The Network State, 250.

17 Srinivasan, 9.

18 Srinivasan, 221.

19 Srinivasan, 221; Craib, Adventure Capitalism, 7.

20 Srinivasan, 245.

21 Srinivasan, 222.

22 Srinivasan, 225.

23 Srinivasan, 233.

24 Srinivasan, 12.

25 Srinivasan, 231.

26 Srinivasan, 26.

27 Craib, Adventure Capitalism, 8; MacDonald, The Political Economy of Non-Territorial Exit, 116.

28 Smith and Burrows, 'Software, Sovereignty and the Post-Neoliberal Politics of Exit', 149.

29 Srinivasan, The Network State, 77.

30 Srinivasan, 236.

31 Srinivasan, 234.

32 Leslie Green, 'Rights of Exit', Legal Theory 4 (1998): 165, 169, https://digitalcommons.osgoode.yorku.ca/scholarly_works/903/.

33 Srinivasan, The Network State, 156.

34 Aihwa Ong, Neoliberalism as Exception: Mutations in Citizenship and Sovereignty (Durham, NC: Duke University Press, 2006).

35 Srinivasan, The Network State, 233.

36 John Locke, A Letter Concerning Toleration, ed. James Tulley (Indianapolis, IN: Hackett Publishing Company, 1983)

37 Craib, Adventure Capitalism, 212-233.

38 Rob Kitchin and Martin Dodge, Code/Space: Software and Everyday Life (Cambridge, MA: MIT Press, 2011), 16.

39 Amanda Greene, 'Consent and political legitimacy', in Oxford Studies in Political Philosophy, vol. 2, eds. David Sobel, Peter Vallentyne and Steven Wall (Oxford: OUP, 2016).

40 Green, Rights of Exit, 175.

41 MacDonald, The Political Economy of Non-Territorial Exit, 123.

42 Margie Cheesman, 'Self-Sovereignty for Refugees? The Contested Horizons of Digital Identity', Geopolitics, 27, no. 1 (2022): 134, 135, DOI: 10.1080/14650045.2020.1823836

43 Green, Rights of Exit, 172.

44 Zajácz, Silk Road, 28.

45 Green, 171, 176.

46 Friedman, Power Without Knowledge, 321.

47 Friedman, 328.

48 Friedman, 322.

49 Craib, Adventure Capitalism, 246.

50 The Blockchain Socialist, Interview with Primavera de Filippi, Overthrowing The Network State: An Initial Critique and Alternatives, podcast audio, January 29, 2023, https://theblockchainsocialist.com/overthrowing-the-network-state-an-initial-critique-and-alternatives/

51 Smith and Burrows, Software, Sovereignty and the Post-Neoliberal Politics of Exit, 149.

52 H.L.A. Hart, The Concept of Law, 3rd ed. (Oxford: OUP, 2012).

53 Green, Rights of Exit, 168; on the possibility of exit penalties in Ayn Rand's capitalist-utopia Galt's Gulch, see Alan Clardy, 'Ayn Rand's Utopian Delusion', Utopian Studies 23, no. 1 (2012): 238, https://www.jstor.org/stable/10.5325/utopianstudies.23.1.0238.

54 Friedman, Power Without Knowledge, 333.

55 Smith and Burrows, Software, Sovereignty and the Post-Neoliberal Politics of Exit, 151.

56 Morshed Mannan and Nathan Schneider, 'Exit to Community: Strategies for Multi-Stakeholder Ownership in the Platform Economy', Georgetown Law Technology Review, 5, no. 1 (2021): 1, https://georgetownlawtechreview.org/exit-to-community-strategies-for-multi-stakeholder-ownership-in-the-platform-economy/GLTR-05-2021/.

57 John Braithwaite, Hilary Charlesworth and Adérito Soares, Networked Governance of Freedom and Tyranny: Peace in Timor-Leste (Canberra: ANU Press, 2012).

58 Primavera de Filippi, "From Competition to Cooperation: Blockchain as a Catalyst for coordiNATION," filmed January 13, 2023 at Plurality Research Network Conference, Berkeley, MA, video, https://www.youtube.com/live/RyObAHv777E?feature=share&t=25557.

Platform Authority and Data Quality: Who Decides What Counts in Data Production for Artificial Intelligence?

Julian Posada
Yale University*

I. Platform Uses and Abuses

Recent applications of OpenAI's Generative Pre-trained Transformer (GPT) family of large language models that predict word sequences have sparked debate on the potential uses and abuses of a technology that produces plausible text on demand. The technology has been tested academically[1] and via court judgments,[2] and questions have been raised about the content and quality of the artificial agent's output and the dangers of reproducing societal harm, as technologies such as facial recognition have done in the past.

In developing its products, the technology industry follows a "big is better" model, thinking that the more data models are trained on, the more accurate their output will be. However, additional training for data models also drives increased computational and human resources.[3] The latest of OpenAI's language models, GPT-4, and its applications, such as ChatGPT, use billions of parameters.[4] These data were scraped from easily accessible internet sites and platforms, but this approach raises concerns over potentially harmful content in the source material.

OpenAI has faces the same problem as many other companies producing data-intensive technologies. As the current paradigm of artificial intelligence (AI) requires large amounts of data, it is unclear how companies can ensure the quality for their products. In other words, how can they prevent their technology from reproducing societal harms? Many companies have hired workers to address this issue. Social-media platforms such as Facebook have hired workers to take down specific posts, and more recent AI companies have hired them to generate and transform their datasets. Many companies see outsourcing labor to lower-income countries as a cost-effective solution that reduces production costs. However, in doing so, they are once again prioritizing data quantity over its quality.

I focus on the relationship between labor and data quality, especially in instances where the generation, annotation, and verification of data is outsourced through digital platforms. My main argument is that higher quality data requires better working conditions because engaged employees whose labor rights are respected provide feedback for improving data quality. In the first part of this paper, I explain the significance of labor in producing data for AI. Then, I will discuss how the industry conceives of the epistemic problem of data production and how the power imbalances in platform labor play a role in this process. The final part of the paper presents recommendations for circumventing epistemic authoritarianism in data production and increasing the quality of data produced through labor outsourcing.

II. Outsourced Labor as the Hidden Ingredient for Artificial Intelligence

In 2005, Amazon launched Mechanical Turk, the first major outsourcing platform, as a form of "artificial AI" intended to distribute tasks related to data production.[5] Its name comes from an 18th-century automaton that seemed capable of playing chess but was in fact controlled by a human concealed inside.[6]

This platform, and those that came later, responds to the need of AI techniques such as machine learning for data and evaluation. Supervised learning requires labeled data, and reinforcement learning requires evaluation. The technology industry, therefore, relies on humans to provide data, annotate it, and verify algorithm outputs.[7]

The need for contemporary technology companies to reduce production costs pushes many to rely on business process outsourcing (BPO) companies or digital platforms for their data work.[8] The former is not the focus of this paper but is worth mentioning. BPO companies are popular mainly for content moderation and algorithmic-verification tasks, and they provide physical infrastructure and workspaces for their employed data workers. One example is the company Sama (previously named Samasource), which employs workers in Kenya and counts OpenAI among its clients.[9] Data-production BPO companies are located all over the world, including India and the Philippines,[10] Argentina and Bulgaria,[11] and the US.[12]

Platforms are primarily headquartered in countries with advanced economies but hire workers from around the world and specialize in different aspects of data production. Some are internal to major technology companies. In addition to Amazon's Mechanical Turk, there are Google's Raterhub and Microsoft's Universal Human Relevance System. Other major players in this ecosystem include Australia's Appen, Canada's Telus.ai, Germany's Clickworker, and the US-based Scale, which recruit workers and provide them with annotation tasks from other platforms, such as those offered by Google and Microsoft.

Workers on these platforms perform four types of tasks, as originally described by Tubaro et al.[13] and detailed by Miceli and Posada[14] in our analysis of over 280 task instructions received by workers. First, platforms provide data *generation*, where workers perform tasks ranging from inputting data to capturing photos, videos, and sound recordings of their surroundings. For example, workers can be tasked with taking photos of themselves in certain positions to train an algorithm to identify them. Second, platforms provide data *annotation*, where workers categorize and give meaning to data. One common task workers perform for autonomous vehicles, for example, is to identify bodies, such as pedestrians, buildings, and other vehicles, that can be encountered while driving. Third, platforms provide *evaluation* of algorithmic outputs by, for example, having workers moderate data for ChatGPT.[15] Fourth, platforms provide *impersonation* of artificial agents, as in the observed case of a worker impersonating a chatbot for a major social-media company.

The persistence of labor in creating and regulating autonomous agents poses several questions. First, do the many workers involved in the data-production pipeline work under decent conditions? Second, are there high standards of data security and privacy, particularly when data are transferred among different users globally? Third, does the data-production process prevent the propagation of harmful content, especially when data generation, annotation, and verification create

meanings or ways of seeing the world that will later be distributed by algorithms? In the next section, I will explore the interrelations between labor rights, data-security standards, and harm avoidance, arguing that companies cannot achieve high-quality data without paying close attention to the social processes involved in producing data.

III. The Ground Truth Problem and Platform Power

When reading dozens of instruction documents for data work, my research team and I were struck by how many of them included managerial elements constantly reminding the workforce that, if they did not perform the tasks according to the clients' design, they would be banned or expelled from the project. In interviews, workers led us to realize that, though most of the tasks were easy and straightforward, some generated disagreement. For example, when moderating social media posts, a Latin American worker disagreed that anti-immigration rhetoric advocating the removal of all immigrants from the US should be protected under freedom of speech. We also encountered complex issues, such as determining the boundaries for "adult content" and issues related to the racial and sexual classification of humans. Thus far, companies can largely self-govern. Although they follow some legal and commercial guidelines, classification decisions are almost left entirely to their discretion.

In managing workers algorithmically and including threats in their instructions, platforms aim at reducing worker bias, a particular problem the data-production industry faces when distributing tasks around globally. Managerial algorithms ensure that results do not deviate from what clients consider correct information or *ground truth*, a term used in the computation and information field. However, as we observed, data classification, especially human and social classification, is subjective and potentially contestable and harmful. Even seemingly straightforward classifications, such as a person crossing the street in an image for autonomous vehicles, can have different labels. Companies tend to classify them as "pedestrians," but such a generic label could preclude manufacturers and their vehicles from considering the particular needs of people who could be labeled as a "child," "person in a wheelchair," or "elderly person."

Managerial algorithms try to reduce worker bias by imposing specific conceptions of ground truth; they also try to reduce risks arising from alienating workers and discouraging feedback. One key difference between smaller data production in BPO companies, which generate data with in-house labelers, and larger platforms, which generate data with freelancers, is that BPO worker engagement and feedback reduce errors and improves data quality and security.

A recent article in *MIT Technology Review* reported that workers in Venezuela hired by US-based Scale AI leaked photos of individuals in private settings, such as in their home bathrooms, taken by development versions of the Roomba, iRobot's robot vacuum cleaner.[16] The company equips these robots with a camera for visualizing their surroundings, and photos from test sites were sent to data workers so they could label the objects in houses. I documented a potentially related episode in which workers were not told what the images they were working on were for and flocked to unmoderated forums on social media to denounce and comment on potential privacy concerns without risking retaliation from their employers.[17]

The fear of being "deactivated" or "banned," terms many platforms use for dismissing a freelance worker, is constant among the dozens of workers we interviewed in Latin America. As platforms' operations are largely unregulated due to the international nature of their transactions, governments and clients do not necessarily compel them to abide by labor laws and regulations. Workers, who are usually located in low-income countries and paid a few cents per task, can be fired without recourse or explanation. Some platforms, such as Australian-based Appen and Canadian-based Telus.ai, are starting to implement contracts with some workers, but this practice is far from being the norm in the sector.

My research project on the platformization of data production has led me to conclude that labor rights and data quality are interrelated. By *quality*, I mean the capacity of data to yield insightful actionable outcomes without reproducing societal harms. Thus far, the industry has relied on underpaid and exploited workers to cost effectively produce data. Companies then utilize managerial algorithms to reduce this alienated workforce's "bias," but those algorithms reproduce the "ground truth" (i.e., the bias) of particular clients and risk the security of the data, which can carry sensitive personal information. In the next section, I present some actionable recommendations to the issue of platform authority.

IV. Recommendations for Data Quality

Advanced AI systems are continuing to perpetuate biases and societal harms, which makes the quest for high-quality data urgent. High-quality data can be achieved in many ways, but here I will underline three methods that relate to data work: ensuring fair-work principles are respected throughout the data-production pipeline; engaging a variety of voices, including those of workers, in AI development; and supporting worker-oriented enterprises.

1. "Labor is not a commodity" is the founding principle of the UN's International Labour Organization. Yet, the rise of the gig economy has enabled the unparalleled commodification of work across myriad sectors, including data production. As decent work is one of the UN's development goals, data-based technologies cannot continue to rely on precarious workforces. Thus, AI developers, platform companies, and regulators should *ensure fair-work principles are respected* throughout the data-production process. The Fairwork Project, a research-oriented initiative from the University of Oxford inspired by the Fairtrade movement, has evaluated different labor platforms across the globe according to the five principles of fair pay, conditions, contracts, management, and representation.[18] To date, none of the platforms evaluated has achieved a perfect score, meaning none of them implements the minimum standards for working conditions. Building upon Clark and Hadfield's concept of regulatory markets for AI, where independent expert institutions inform the public of compliance with regulations,[19] I argue a thorough evaluation of data-production platforms, either through government action or independent research, could elucidate the working conditions across the sector and inform different stakeholders, including AI companies, and thereby potentially induce a race to the top and thus compliance with labor rights and laws.

2. Data quality is also a question of governance. The case of outsourced data production links discussions on data, platforms, and AI governance. For

example, digital platforms have created internal governance mechanisms such as Meta's Oversight Board for content moderation policies. External governance has also come in the form of regulations such as the General Protection Data Regulation. Moreover, co-governance mechanisms, such as the Global Network Initiative and the Partnership on AI, are examples of third-party entities that can steer policies governing data-based goods and institutions.[20] I recommend governance mechanisms that *enable worker input in the data-production pipeline*. Miceli and Posada's research has shown that worker input and feedback on tasks in BPO settings are crucial in improving data quality.[21] Data workers have expertise and unique perspectives because they handle data directly. Their insights could prove crucial to identifying errors in generation, classification, and verification processes.

3. Ethical AI cannot exist without ethical data-production processes that guarantee worker well-being. However, endeavors that guarantee working standards are difficult to conceive and operationalize due to a lack of labor standards in the data-production sector and the race to the bottom that characterizes digital labor outsourcing because clients expect access to data at lower costs. Several impact source initiatives, such as CloudFactory, iMerit, and Sama, have emerged in the data-production sector in recent years. However, as Kaye stresses, the lack of governance mechanisms and involvement from civil society in these issues renders the proliferation and accountability of such ethical endeavors difficult.[22] Even supposedly ethical organizations have been criticized for their labor practices. For example, a recent *Time* article documented possible union-busting practices from Sama.[23] Standards and mechanisms of accountability should be created while *supporting worker-centered initiatives,* including impact-sourcing companies, cooperatives, and not-for-profits, that respect the standards mentioned above to mitigate the race-to-the-bottom trend in platform labor.

In this paper, I have described the importance of labor in the production of data and the subsequent development of data-based technologies such as AI. The current system is one of self-governance and increasing platform authority, where profits are prioritized over high-quality data, that is, data that produce insightful outcomes without reproducing societal harms. There cannot be high-quality data and ethical AI systems without respect for human rights—including labor rights. Therefore, the industry should strive to respect fair-work principles, enable worker feedback in the data-production pipeline, and support worker-centered initiatives backed by standards and effective governance. These initiatives allow for the broader considerations necessary to democratize digital spaces and entities, including platforms, and reduce power concentration among a few entities.

Endnotes

* Acknowledgement and disclaimer: The views and positions expressed in this report are solely those of the author and do not necessarily reflect the views of the Department of Foreign Affairs, Trade and Development (commonly known as Global Affairs Canada) or the Government of Canada. The report is in its original language.

1 Chris Westfall, "Educators Battle Plagiarism As 89% Of Students Admit To Using OpenAI's ChatGPT For Homework," Forbes, 2023, https://www.forbes.com/sites/chriswestfall/2023/01/28/educators-battle-plagiarism-as-89-of-students-admit-to-using-open-ais-chatgpt-for-homework/.

2 Luke Taylor, "Colombian Judge Says He Used ChatGPT in Ruling," The Guardian, February 3, 2023, https://www.theguardian.com/technology/2023/feb/03/colombia-judge-chatgpt-ruling.

3 Emily M. Bender et al., "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? ,," in Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (ACM, 2021), 610–23, https://doi.org/10.1145/3442188.3445922.

4 Open AI, "GPT-4 Technical Report," 2023, https://cdn.openai.com/papers/gpt-4.pdf

5 Mary L. Gray and Siddharth Suri, Ghost Work: How to Stop Silicon Valley from Building a New Global Underclass (Houghton Mifflin Harcourt, 2019).

6 Antonio A. Casilli, En Attendant Les Robots: Enquête Sur Le Travail Du Clic, La Couleur Des Idées (Éditions du Seuil, 2019).

7 Paola Tubaro, Antonio A Casilli, and Marion Coville, "The Trainer, the Verifier, the Imitator: Three Ways in Which Human Platform Workers Support Artificial Intelligence," Big Data & Society 7, no. 1 (January 2020): 205395172091977, https://doi.org/10.1177/2053951720919776; Milagros Miceli and Julian Posada, "The Data-Production Dispositif," Proceedings of the ACM on Human-Computer Interaction 6, no. CSCW2 (2022).

8 Antonio A. Casilli and Julian Posada, "The Platformisation of Labor and Society," in Society and the Internet: How Networks of Information and Communication Are Changing Our Lives, ed. Mark Graham and William H. Dutton, Vol. 2 (Oxford University Press, 2019).

9 Billy Perrigo, "Inside Facebook's African Sweatshop," Time, February 14, 2022, https://time.com/6147458/facebook-africa-content-moderation-employee-treatment/.

10 Sana Ahmad and Martin Krzywdzinski, "Moderating in Obscurity: How Indian Content Moderators Work in Global Content Moderation Value Chains," in Digital Work in the Planetary Market, ed. Mark Graham and Fabian Ferrari (The MIT Press, 2022), https://doi.org/10.7551/mitpress/13835.001.0001.

11 Milagros Miceli et al., "Documenting Data Production Processes: A Participatory Approach for Data Work," Proceedings of the ACM on Human-Computer Interaction 6, no. CSCW2 (2022).

12 Sarah T. Roberts, Behind the Screen: Content Moderation in the Shadows of Social Media ; with a New Preface (Yale University Press, 2021).

13 Tubaro, Casilli, and Coville, "The Trainer, the Verifier, the Imitator."

14 Miceli and Posada, "The Data-Production Dispositif."

15 Billy Perrigo, "Exclusive: The $2 Per Hour Workers Who Made ChatGPT Safer," Time, January 18, 2023, https://time.com/6247678/openai-chatgpt-kenya-workers/.

16 Eileen Guo, "A Roomba Recorded a Woman on the Toilet. How Did Screenshots End up on Facebook?," MIT Technology Review, accessed February 22, 2023, https://www.technologyreview.com/2022/12/19/1065306/roomba-irobot-robot-vacuums-artificial-intelligence-training-data-privacy/.

17 Julian Posada, "Family Units. The Communities behind the Data Annotation Work That Powers AI," Logic Magazine, 2021, https://logicmag.io/beacons/family-units/.

18 Mark Graham et al., "The Fairwork Foundation: Strategies for Improving Platform Work in a Global Context," Geoforum 112 (2020): 100–103, https://doi.org/10.1016/j.geoforum.2020.01.023.

19 Jack Clark and Gillian K. Hadfield, "Regulatory Markets for AI Safety" (arXiv, 2019), http://arxiv.org/abs/2001.00078.

20 Robert Gorwa, "What Is Platform Governance?," Information, Communication & Society 22, no. 6 (2019): 854–71, https://doi.org/10.1080/1369118X.2019.1573914.

21 Miceli and Posada, "The Data-Production Dispositif."

22 D. Kaye, "A Human Rights Approach to Platform Content Regulation," Report of the UN Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, 2018, https://freedex.org/a-human-rights-approach-to-platform-content-regulation/.

23 Perrigo, "Inside Facebook's African Sweatshop."