Centre for International
Governance Innovation

Policy Brief No. 187 — July 2024

How Authoritarian Value Systems Undermine Global AI Governance

Sabhanaz Rashid Diya

Key Points

- Digital authoritarianism, including in artificial intelligence (AI) technologies, is often considered an issue limited to a few illiberal regimes. However, neo-liberal AI technologies can be equally pervasive. It is crucial to treat authoritarianism as a *values* complex that permeates both autocratic and liberal societies.
- Authoritarian values may manifest through the transplant of legal practices between states, autocratic homogenization through multilateral mechanisms, and the exploitation of geopolitical tensions to adopt protectionist policies. These approaches exacerbate public polarization around AI governance by creating a false dichotomy between innovation and sovereignty on the one hand, and fundamental rights on the other, chipping away at institutional trust.
- Policy solutions to mitigate the erosion of democratic norms and public trust should focus on international mechanisms central to AI governance. These mechanisms need to introduce procedural safeguards that ensure transparency and accountability through equitable multi-stakeholder processes. Additionally, they should encourage regulatory diversity tailored to sociopolitical contexts and aligned with international human rights principles.

Introduction

In recent years, the rapid advancement of artificial intelligence (AI) technologies has prompted a global push to regulate the AI industry. As AI has become increasingly integrated into various aspects of society, from health care and education to communication, concerns about its potential risks and implications have also grown. Governments, international organizations and rights-based advocacy groups are grappling with the need to develop comprehensive regulatory frameworks that balance innovation and progress with ethical considerations.

At the heart of the AI regulation debate is its impact on trust, particularly in safeguarding human dignity and agency while offering immense potential to advance human lives and socio-economic well-being. Trust depends on predictability (Hardin 2002) and shared normative values (Lahno 2001). The "black-box problem" of AI systems, characterized by the lack of transparency in human decisions behind these systems, poses significant limitations in bridging trust (von Eschenbach 2021). "Black-box systems" refer to deep learning algorithms that have complex structures and learning models, generating results that may not be intrinsically interpretable to humans (Hall and Gill 2019). In the absence of a clear explanation for why an AI system behaves in a certain way, making it unpredictable, confusion arises about what specifically needs to be governed to prevent harmful outcomes. Moreover, some

About the Author

Sabhanaz Rashid Diya is a CIGI senior fellow and the founder of Tech Global Institute, a global tech policy non-profit focused on advancing equity in design and governance of technologies in the global majority. She has advised governments in 20 countries, including leading closed-door briefings with the White House, multilateral international organizations and bilateral donors on global internet and platform governance, responsible artificial intelligence (AI) and human rights.

A computational social scientist by training, Sabhanaz has more than 17 years of experience at the intersection of technology policy, ethics and international development. She was most recently the head of public policy for Bangladesh at Meta, where she led on various regulatory and legislative issues, including privacy, online harms and algorithmic transparency. Sabhanaz also worked at the Bill & Melinda Gates Foundation, leading policy and advocacy efforts in digital identity, data governance and AI. Her career spans the private and public sectors in the United States, Asia and Africa on encryption policy, digital trade, AI applications in the global majority and internet governance. She is a visiting policy fellow at the Oxford Internet Institute, and a member of the Advisory Network of the Freedom Online Coalition.

AI systems inherently risk undermining normative values. One of the most widely known examples of value violation is the use of the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) technology in the United States' court system to "predict" recidivism for over one million defendants since its release in the early 2000s. An investigation on COMPAS found that Black defendants were twice as likely as their white counterparts to be, incorrectly, judged to be at a higher risk of recidivism (Larson et al. 2016).

Citizens expect their elected public officials to mitigate the negative impacts of AI systems. For example, an overwhelming majority of Americans surveyed want AI regulation, with 67 percent indicating their concern that the government will not go *far enough* in regulating their use (Faverio and Tyson 2023). Ninety-one percent of surveyed Europeans support state-led "careful management of AI" (Dreksler et al. 2023). Although opinion surveys in recent years have documented a global decline in trust in government, there is a general consensus that democratically elected policy makers are trustworthy brokers of the public's interest.

Conversely, according to researchers at the V-Dem (Varieties of Democracy) Institute at the University of Gothenburg, 72 percent of the global population reside in an autocracy — a significant proportion of which are in low- and middle-income regions (Papada et al. 2023). *The Economist*'s annual Democracy Index similarly finds that only eight percent of the world's population live in a "full democracy," predominantly in North America, Western Europe and Australia (Economist Intelligence Unit 2024). This raises the critical question of whether authoritarian actors can be *trusted* to develop AI regulations in the public interest.

This policy brief argues that authoritarianism constitutes a value complex present in both autocratic and democratic societies, and these values can be transmitted through any technology that is developed in these environments. As illustrated by COMPAS, Western neo-liberal AI systems pose just as high a risk of eroding trust as those developed under purely autocratic conditions. It is critical for policy makers to evaluate digital authoritarianism holistically through the lens of the values influencing both the deployment of AI systems *and* their governance structures globally, rather than by focusing solely on the impact of a few regimes.

AI and Digital Authoritarianism

Existing literature on digital authoritarianism, particularly concerning AI systems, skews heavily toward Chinese and Russian technologies and the autocratic regimes behind them. China has implemented a comprehensive domestic surveillance architecture with more than 200 million cameras, enabling widespread facial recognition for crime prevention and sophisticated profiling (Mozur 2018). A survey by Beijing News Think Tank found that nearly 80 percent of Chinese respondents oppose automatic facial recognition in commercial zones in Beijing, and 96 percent are concerned about privacy and data breach (as reported by Masha Borak [2021]). For its part, Russia has invested in AI-enabled military warfare and automated influence operations. Despite doubling its defence spending in 2023 to more than US$100 billion, Russia's defence ministry only allocates an estimated US$12 to US$36 million annually for AI research — significantly less than China's projected US$150 billion by 2030 (Petrella, Miller and Cooper 2021).

The international security discourse has principally focused on Russia's and China's export of authoritarian technologies to other regions (Feldstein 2019a). In recent years, concern has been growing about the unregulated distribution of neo-liberal AI technologies developed under democratic conditions, leading to surveillance capitalism transforming into the surveillance state (Zuboff 2019). The use of black-box algorithms in curating and microtargeting information has contributed to regime stability globally and exacerbated Orwellian efforts (Gunitsky 2015). Among 176 countries, 75 have implemented AI surveillance technologies; 32 countries are using surveillance technology from US firms (Feldstein 2019b). There is an increasing trend of exporting authoritarian value systems globally — through both technologies and their governing rules — negatively affecting the public's trust in innovation and policy making.

The "Illiberal Values Test" for Global AI Governance

Current policy discourse often assumes that AI applications can be governed against the backdrop of a predefined set of values and legal practices (Nemitz 2018). In reality, the laws and values applied during the design and development of AI systems are often specific — and limited — to the unique environments in which they were created (Gordon, Rieder and Sileno 2022). The practical downstream impacts of AI applications can lead to legal and ethical implications that differ from these applications' theoretical intentions.

The disconnect between intent and application is exemplified by a case in India. In March 2024, the Indian government issued a non-binding advisory requiring AI providers (platforms and intermediaries) to seek explicit permission from the Ministry of Electronics & Information Technology (MeitY) before deploying any "unreliable" or underdeveloped AI models on "Indian Internet" (Kalra and Vengattil 2024). Notably, this move followed an exchange with Google's chatbot Gemini responding to the question "Is Modi a fascist" and allegations of Google's violation of Rule 3(1)(b) of MeitY's *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* (*IT Rules, 2021*).[1]

Governments worldwide, including India's, are putting an emphasis on regulating bias in AI systems; however, these attempts rarely differentiate between conceptual bias (rendering AI systems as "unreliable") and its actual effect on exacerbating discrimination. India's political context demonstrates weaponization of the theoretical premise of bias to stifle rights and expression. In fact, the very reference to the *IT Rules, 2021* as a legal intervention to govern bias raises concerns. These rules have faced significant human rights scrutiny in India, with 17 petitions

1 *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021* (06 April 2023) Ministry of Electronics and Information Technology, online: <www.meity.gov.in/writereaddata/files/Information%20Technology%20%28Intermediary%20Guidelines%20and%20Digital%20Media%20Ethics%20Code%29%20Rules%2C%202021%20%28updated%2006.04.2023%29-.pdf>.

in courts challenging their constitutionality. They are criticized for being overly broad and mandating content removal under vague provisions, such as "information which is patently false or misleading," as well as for breaking encryption and threatening to revoke safe harbour if a regulated intermediary is non-compliant (Global Network Initiative 2023). Applying these rules to AI without resolving existential legal issues not only reflects India's ambitions in the global AI regulatory race, but also highlights that rushed efforts to regulation could potentially risk violating democratic values.

Authoritarian AI policies are typically viewed as a domestic issue, unique to countries with authoritarian and hybrid regimes, where they serve to consolidate state control over private powers. The other end of the spectrum includes legislative interventions such as the European Union's AI Act, which serves to bring democratic control over private powers. European policy makers hope that the AI Act will establish a global benchmark akin to the Brussels Effect seen with the EU General Data Protection Regulation (GDPR) (Bradford 2019; European Parliament 2023). However, should the AI Act trigger a de jure Brussels Effect (ibid.), the question will arise as to its intended impact on democracy, when other nations explicitly model their AI laws after European rules.

To address this question, it is crucial to understand the normative values and institutions underpinning the AI Act. The law operates within a broader human rights framework, encompassing the Council of Europe through the European Convention on Human Rights, the Charter of Fundamental Rights of the European Union, and national constitutions. The powers of the EU AI Office, national competent authorities designated by member states, and law enforcement are bound by well-established, rights-centric legal constraints at both the European Commission and national levels. While these constraints may not fully address the European civil society concerns regarding human rights and the rule of law under the AI Act (Civil Liberties Union for Europe 2023), they provide necessary guardrails, including human rights impact assessment for high-risk AI systems, to mitigate various risks of abuse.

In contrast, two of the world's most populous democracies, India and Brazil, rank seventy-ninth and eighty-third, respectively, on the *Rule of Law Index 2023*, which considers rule of law progress and setbacks across 142 countries and jurisdictions (World Justice Project 2023, 10). These scores indicate a lack of checks and balances on government powers, a convergence of the administrative and judicial bodies, and poor records in safeguarding fundamental rights. If the AI Act were enforced outside the European Union under a different set of normative values and institutions, it would likely fail the democracy test. Without mature, trustworthy institutions, any rules developed under democratic conditions can be exploited under authoritarian rule.

For instance, article 5 of the AI Act prohibits the use in the European Union of AI systems that deploy "purposefully manipulative or deceptive techniques" with the objective or effect of impairing a person's ability to make informed decisions.[2] This categorization prompts questions about the *definition* of "manipulative" techniques, the designated authority's process for determining purposeful manipulation, and the methods used for measuring the impact of that attempted manipulation on a person's decision making. Article 70 grants power to one or more designated "national competent authorities" to enforce the regulation, emphasizing that these powers must be exercised "independently, impartially and without bias."[3]

Under undemocratic circumstances, these provisions allow politically or ideologically vested state authorities to wield their powers to serve specific interests. If the AI Act is reinforced with a de jure Brussels Effect in the absence of embedded institutional and procedural safeguards in India, it could potentially empower the government to prohibit Google Gemini or misuse its information-gathering powers to undermine privacy, information rights and good-faith safe harbour protections. Such activity might be undertaken alongside allegations that Google was deploying deceptive techniques to mislead Indian users.

Similarly, article 22 of the AI Act requires non-EU providers of AI systems to appoint an "authorised representative" established in the Union.[4] In many authoritarian and hybrid regimes, such as those of Brazil, India, Nigeria, Türkiye and Vietnam, physical presence within the nation is tied to

2 See https://artificialintelligenceact.eu/article/5/.

3 See https://artificialintelligenceact.eu/article/70/.

4 See https://artificialintelligenceact.eu/article/22/.

criminal liabilities, also known as hostage laws (Elliott 2021). Countries with a history of cracking down on internet freedoms have used hostage laws to intimidate staff or authorized representatives of technology companies, notably social media platforms, to comply with overly broad government orders. Many countries indicate they were inspired by Germany's Network Enforcement Law (Netzwerkdurchsetzungsgesetz, or NetzDG), which came into effect in 2017, that requires social media companies to appoint a local German representative and swiftly remove illegal content or else pay substantive fines. According to the human rights think tank Justitia, a replication of the NetzDG has served as a template for censorship in 13 countries (Mchangama and Fiss 2019, 17).

Some argue that these comparisons are unfair. Laws are written for specific contexts operating within legal and political institutions in a particular jurisdiction and should not be expected to address concerns elsewhere. The argument lies therein. Laws tailored to specific socio-legal contexts should not be promoted for imitation in other countries or contexts by policy makers and experts.

Trust in AI regulations depends on technology-agnostic external factors, such as political climate and institutional maturity. An innocuous set of rules can easily become a dictator's tool in contexts characterized by authoritarian values.

AI Authoritarianism through Multilateral Mechanisms

Multilateral diplomacy has historically upheld the rule-based liberal world order (Ikenberry 2005; 2015), yet recent years have seen challenges from authoritarian regimes (Ginsburg 2020). Policy discourse has typically framed multilateralism through the lens of liberal multilateralism, overlooking authoritarian multilateralism (Raymond and Sherman 2024). Authoritarian multilateralism prioritizes collective economic rights over individual rights and makes broader allowances for power privileges, leading to systemic violations of citizen human rights (ibid.). Authoritarian actors are reshaping *procedural* rules of international multilateral forums to align cyber and AI policies more consistently with their values, primarily employing three tactics.

First, authoritarian actors use politicized language (Giles and Hagestad 2013) in UN and other multilateral proposals to disguise political desires for information control. For instance, over the past decade, delegations from Beijing and Moscow at the UN General Assembly proposed cyber resolutions using terms such as "information security" and "cybercrime," presumably to advance legitimate domestic policy making. In reality, these terms contorted the *form* of liberal multilateralism by broadening the interpretation of "information security" beyond those of liberal democracies such as Australia, South Korea and the United States, inadvertently validating information *control* (Raymond and Sherman 2024). This process co-opts "liberal" terms in multilateral mechanisms to solidify authoritarian regime stability, contributing to shifts "both to an international order which is less liberal, as well as to the global weakening of liberal social and political norms within states" (Bettiza and Lewis 2020, 571).

Second, authoritarian actors incorporate collectivist language on societal and economic welfare to invest in funding schemes for state-led technology control and to gain legitimacy through multilateral mechanisms. Saudi Arabia, for example, is planning a US$40-billion AI fund in collaboration with Silicon Valley venture capitalists (Farrell and Copeland 2024) through its Public Investment Fund (PIF), a sovereign wealth fund of more than US$900 billion with Prince Mohammed bin Salman as its chairperson.[5] PIF's portfolio includes major investments in companies that include India's Jio Platforms, as well as bilateral funding pools, such as the Russian Direct Investment Fund (PIF 2017) and Japan's SoftBank (PIF 2023). Since 2016, Saudi Arabia and China have expanded bilateral technoscientific cooperation, with 60 percent of joint projects undertaken by Chinese state-owned enterprises, furthering both nations' political objectives (Al-Sudairi, Hai and Alahmad 2023).

Saudi Arabia's foreign policy is centred on its technological ambitions, which it promotes through international organizations (ibid.). In 2017, it established the Digital Cooperation Organization to coordinate technology cooperation among Global South countries and endorsed the Riyadh AI Call

5 See www.pif.gov.sa/en/who-we-are/our-leadership; PIF (2023).

for Action Declaration. In 2020, Saudi Arabia joined the UN Commission on Science and Technology for Development (CSTD) and was elected to chair its twenty-fifth session (Arab News 2022). Despite strong objections from more than 100 human rights groups (Access Now 2023), Saudi Arabia is set to host this year's Internet Governance Forum, amid allegations of using spyware to surveil and censor its citizens and journalists and illegally accessing personal information (ibid.).

Third, systemic disparities in power politics between the Global North and the Global South prompt collaborations between countries in passing resolutions — often based on authoritarian value systems — through multilateral processes to oppose Western hegemony. In November 2019, China and Russia co-sponsored a cybercrime resolution in the UN General Assembly's Third Committee with support from Angola, Egypt, India, North Korea, Syria and Uganda that would sideline multi-stakeholderism (Raymond and Sherman 2024). This process of "bad-faith invocation of liberal multilateral principles" (ibid., 124; Pouliot 2021) increasingly diffuses authoritarian value systems between countries, evident through the adoption of repressive protectionist policies.

The Global North-South Divide and AI Governance

Countries once codependent on Western liberal democracies are increasingly frustrated by unfulfilled promises on trade and financial commitments, and inconsistent application of human rights norms during bilateral diplomacy. During the COVID-19 pandemic, "wealthier" countries not only reneged on financial commitments to equitable vaccine distribution, but also secured early vaccine access through direct deals with manufacturers, procuring most of the world's vaccine supply before other countries had access (Suzman 2023). Similarly, Western donors have fallen short of targets by tens of billions of dollars every year since agreeing to an annual $100-billion commitment to support climate adaptation in developing countries at the 2015 Paris climate summit (ibid.).

The widening global North-South rift poses significant challenges for advancing a global consensus on AI governance. While Western democracies advocate for liberal values, their technology remains prohibitively expensive and inaccessible to many developing countries (Unver 2021). Lacking resources like compute power and expensive hardware, these countries turn to more dependable (and recent) allies, such as China and Saudi Arabia, who can support affordable digital infrastructure. Even European "norm superpowers" such as Austria, Hungary and the Netherlands, despite denouncing China-led AI authoritarianism, are planning to acquire Chinese 5G (fifth-generation) mobile networks and infrastructure (ibid.). These discrepancies undermine trust and exacerbate ideological and investment disparities, leading to fragmented AI governance based on competing value systems.

This division is evident in India's approach in global AI initiatives through its leadership at the Global Partnership on Artificial Intelligence (GPAI) and the Group of Twenty (G20). The country's efforts, particularly through the inclusion of the African Union in the G20 and export of the digital public infrastructure, have garnered trust among Global South policy makers, leading to widespread support for the New Delhi "GPAI Ministerial Declaration."[6] By premising its policies on *inclusive* AI development in the Global South, India has skyrocketed as a normative superpower in AI governance, allowing it to exert significant influence over other countries to align with its protectionist political ideology.

6 See https://gpai.ai/2023-GPAI-Ministerial-Declaration.pdf.

Recommendations

The widespread adoption of authoritarian value systems in AI policies worldwide is symptomatic of systemic trust deficits within and between countries. To mitigate risks of "authoritarian takeover," policy makers and experts should promote regulatory resilience through transparency, capacity and global multi-stakeholder alliances.

Promote Transparency and Accountability

Existing multilateral and international forums on AI governance lack transparency, particularly in their process for consulting stakeholders and considering different policy options for a rapidly evolving technology. Both authoritarian capture and regulatory capture are perceived to be prevalent and limit trust in rule-based, democratic governance. There need to be more robust requirements for participating nations and stakeholders to disclose potential conflicts of interest, and independent oversight of the rule-making process to ensure that it serves the public interest.

Strengthen Global Multi-stakeholder Alliances

Liberal democracies are increasingly opposing multilateral modalities of rule-making, and moving toward accommodating non-state actors to counter authoritarian influence. There is an urgent need to invest in building robust and inclusive measures to establish procedural safeguards around *global* multi-stakeholder systems, especially ensuring the inclusion of Global South voices. Existing processes are widely perceived to be ineffective, inequitable and non-representative of the public interest. A careful deliberation on criteria for broadening access, inclusion, negotiation, transparency and short- and long-term accountability will be crucial to rebuilding trust in both institutions and processes.

Promote Regulatory Diversity through Meaningful Investments

Countries operate on different levels of political and institutional maturity; therefore, it is necessary for them to pursue diverse regulatory options while adhering to democratic values. Resource-constrained governments need "norm export" to be supplemented with tangible investments in inclusive, rights-binding digital infrastructure, bringing stronger voices on the international AI policy agenda, and holistic capacity-building initiatives for non-state actors in countries to strengthen accountability.

Conclusion

The global landscape of AI governance is deeply influenced by geopolitical dynamics, particularly the growing rift between the Global North and the Global South and the shift away from Western hegemony. The dichotomy between autocratic and democratic states demands re-evaluation, recognizing that the core issue lies with conflicting value systems rather than with regime types. The increasing export of authoritarian principles — through both technology and regulatory practices — undermines institutional trust and threatens democratic norms. This dynamic raises critical questions about the capacity of various states to develop AI regulations that genuinely serve the public interest.

To mitigate these risks, it is essential to view authoritarianism as a global value complex and focus on strengthening international governance mechanisms. These mechanisms should promote transparency and regulatory diversity tailored to sociopolitical contexts by fostering and empowering not only multi-stakeholder alliances as accountability interventions but also stronger voices from diverse Global South communities in the international AI agenda. Additionally, international investments mechanisms must support rights-binding digital infrastructures that promote policy resiliency through both state and non-state solutions.

Acronyms and Abbreviations

AI	artificial intelligence
COMPAS	Correctional Offender Management Profiling for Alternative Sections
CSTD	Commission on Science and Technology for Development
G20	Group of Twenty
GDPR	General Data Protection Regulation
GPAI	Global Partnership on Artificial Intelligence
MeitY	Ministry of Electronics & Information Technology
NetzDG	Netzwerkdurchsetzungsgesetz (Network Enforcement Law)
PIF	Public Investment Fund
V-Dem	Varieties of Democracy

Works Cited

Access Now. 2023. "Joint Statement: Internet Governance Forum must reverse decision to make Saudi Arabia its next host." Press release, October 12. www.accessnow.org/campaign/igf-reverse-saudi-arabia-host-decision/.

Al-Sudairi, Mohammed, Steven Jiawei Hai and Kameal Alahmad. 2023. "How Saudi Arabia Bent China to Its Technoscientific Ambitions." Carnegie Endowment for International Peace, August 1. https://carnegieendowment.org/research/2023/08/how-saudi-arabia-bent-china-to-its-technoscientific-ambitions?lang=en.

Arab News. 2022. "Saudi Arabia chairs 25th session of Commission on Science and Technology for Development." April 10. https://arab.news/n653g.

Bettiza, Gregorio and David Lewis. 2020. "Authoritarian Powers and Norm Contestation in the Liberal International Order: Theorizing the Power Politics of Ideas and Identity." *Journal of Global Security Studies* 5 (4): 559–77. https://doi.org/10.1093/jogss/ogz075.

Borak, Masha. 2021. "Facial recognition is used in China for everything from refuse collection to toilet roll dispensers and its citizens are growing increasingly alarmed, survey shows." *South China Morning Post*, January 27. www.scmp.com/tech/innovation/article/3119281/facial-recognition-used-china-everything-refuse-collection-toilet.

Bradford, Anu. 2019. *The Brussels Effect: How the European Union Rules the World*. Oxford, UK: Oxford University Press.

Civil Liberties Union for Europe. 2023. "The AI Act Must Protect the Rule of Law." Open letter, September 28. www.liberties.eu/en/stories/ai-act-rule-of-law/44917.

Dreksler, Noemi, David McCaffary, Lauren Kahn, Kate Mays, Markus Anderljung, Allan Dafoe, Michael C. Horowitz and Baobao Zhang. 2023. "Preliminary Survey Results: US and European Publics Overwhelmingly and Increasingly Agree That AI Needs to Be Managed Carefully." Centre for Governance of AI, April 17. www.governance.ai/post/increasing-consensus-ai-requires-careful-management.

Economist Intelligence Unit. 2024. "Democracy Index: conflict and polarisation drive a new low for global democracy." News release, February 15. www.eiu.com/n/democracy-index-conflict-and-polarisation-drive-a-new-low-for-global-democracy/.

Elliott, Vittoria. 2021. "New laws requiring social media platforms to hire local staff could endanger employees." *Rest of World*, May 14. www.restofworld.org/2021/social-media-laws-twitter-facebook/.

European Parliament. 2023. "Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI." Press release, December 9. www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai.

Farrell, Maureen and Rob Copeland. 2024. "Saudi Arabia Plans $40 Billion Push Into Artificial Intelligence." *The New York Times*, March 24. www.nytimes.com/2024/03/19/business/saudi-arabia-investment-artificial-intelligence.html.

Faverio, Michelle and Alec Tyson. 2023. "What the data says about Americans' views of artificial intelligence." Pew Research Center, November 21. www.pewresearch.org/short-reads/2023/11/21/what-the-data-says-about-americans-views-of-artificial-intelligence/.

Feldstein, Steven. 2019a. "The Road to Digital Unfreedom: How Artificial Intelligence is Reshaping Repression." *Journal of Democracy* 30 (1): 40–52. https://doi.org/10.1353/jod.2019.0003.

———. 2019b. "The Global Expansion of AI Surveillance." Carnegie Endowment for International Peace, September 17. www.jstor.org/stable/resrep20995.1.

Giles, Keir and William Hagestad II. 2013. "Divided by a Common Language: Cyber Definitions in Chinese, Russian and English." In "Chapter 5. Cyber Conflict — Politics, Semantics, Ethics and Moral," *5th International Conference on Cyber Conflict: Proceedings*, edited by K. Podins, J. Stinissen and M. Maybaum, 413–30. Talinn, Estonia: NATO Cooperative Cyber Defence Centre of Excellence. https://ccdcoe.org/uploads/2018/10/CyCon_2013_Proceedings.pdf.

Ginsburg, Tom. 2020. "How Authoritarians Use International Law." *Journal of Democracy* 31 (4): 44–58. www.journalofdemocracy.org/articles/how-authoritarians-use-international-law/.

Global Network Initiative. 2023. "GNI Analysis: Information Technology Rules Put Rights at Risk in India." Press release, March 7. www.globalnetworkinitiative.org/india-it-rules-2021.

Gordon, Geoff, Bernhard Rieder and Giovanni Sileno. 2022. "On mapping values in AI governance." *Computer Law & Security Review* 46: 105712. https://doi.org/10.1016/j.clsr.2022.105712.

Gunitsky, Seva. 2015. "Corrupting the Cyber-Commons: Social Media as a Tool of Autocratic Stability." *Perspectives on Politics* 13 (1): 42–54. https://doi.org/10.1017/S1537592714003120.

Hall, Patrick and Navdeep Gill. 2018. *An Introduction to Machine Learning Interpretability*. 2nd ed. Sebastopol, CA: O'Reilly Media.

Hardin, Russell. 2002. *Trust and Trustworthiness*. New York, NY: Russell Sage Foundation.

Ikenberry, G. John. 2005. "Power and liberal order: America's postwar world order in transition." *International Relations of the Asia-Pacific* 5 (2): 133–52. https://doi.org/10.1093/irap/lci112.

———. 2015. "The Future of Multilateralism: Governing the World in a Post-Hegemonic Era." *Japanese Journal of Political Science* 16 (3): 399–413. https://doi.org/10.1017/S1468109915000158.

Kalra, Aditya and Munsif Vengattil. 2024. "India asks tech firms to seek approval before releasing 'unreliable' AI tools." Reuters, March 4. www.reuters.com/world/india/india-asks-tech-firms-seek-approval-before-releasing-unreliable-ai-tools-2024-03-04/.

Lahno, Bernd. 2001. "On the Emotional Character of Trust." *Ethical Theory and Moral Practice* 4: 171–89. https://doi.org/10.1023/A:1011425102875.

Larson, Jess, Surya Mattu, Lauren Kirchner and Julia Angwin. 2016. "How We Analyzed the COMPAS Recidivism Algorithm." ProPublica, May 23. www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

Mchangama, Jacob and Joelle Fiss. 2019. "The Digital Berlin Wall: How Germany (Accidentally) Created a Prototype for Global Online Censorship." Justitia, November 5. https://justitia-int.org/wp-content/uploads/2019/11/Analyse_The-Digital-Berlin-Wall-How-Germany-Accidentally-Created-a-Prototype-for-Global-Online-Censorship.pdf.

Mozur, Paul. 2018. "Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras." *The New York Times*, July 8. www.nytimes.com/2018/07/08/business/china-surveillance-technology.html.

Nemitz, Paul. 2018. "Constitutional democracy and technology in the age of artificial intelligence." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376 (2133): 20180089. https://doi.org/10.1098/rsta.2018.0089.

Papada, Evie, David Altman, Fabio Angiolillo, Lisa Gastaldi, Tamara Köhler, Martin Lundstedt, Natalia Natsika, et al. 2023. *Defiance in the Face of Autocratization: Democracy Report 2023*. Gothenburg, Sweden: V-Dem Institute. www.v-dem.net/documents/29/V-dem_democracyreport2023_lowres.pdf.

Petrella, Stephanie, Chris Miller and Benjamin Cooper. 2021. "Russia's Artificial Intelligence Strategy: The Role of State-Owned Firms." *Orbis* 65 (1): 75–100. https://doi.org/10.1016/j.orbis.2020.11.004.

PIF. 2017. "RDIF strengthens cooperation with PIF of Saudi Arabia." Press release, September 27. www.pif.gov.sa/en/news-and-insights/press-releases/2017/rdif-strengthens-cooperation/.

———. 2023. *Shaping the Future: Annual Report 2022*. Riyadh, Saudia Arabia: PIF. www.pif.gov.sa/en/our-financials/annual-reports/.

Pouliot, Vincent. 2021. "The Gray Area of Institutional Change: How the Security Council Transforms Its Practices on the Fly." *Journal of Global Security Studies* 6 (3): ogaa043. https://doi.org/10.1093/jogss/ogaa043.

Raymond, Mark and Justin Sherman. 2024. "Authoritarian multilateralism in the global cyber regime complex: The double transformation of an international diplomatic practice." *Contemporary Security Policy* 45 (1): 110–40. https://doi.org/10.1080/13523260.2023.2269809.

Suzman, Mark. 2023. "The Roots of the Global South's New Resentment: How Rich Countries' Selfish Pandemic Responses Stoked Distrust." *Foreign Affairs*, September 8. www.foreignaffairs.com/africa/roots-global-souths-new-resentment.

Unver, Akin. 2021. "Motivations for the Adoption and Use of Authoritarian AI Technology." In *Issues on the Frontlines of Technology and Politics*, edited by Steven Feldstein, 15–16. Washington, DC: Carnegie Endowment for International Peace. https://carnegie-production-assets.s3.amazonaws.com/static/files/202110-Feldstein_Frontlines_final3.pdf.

von Eschenbach, Warren J. 2021. "Transparency and the Black Box Problem: Why We Do Not Trust AI." *Philosophy & Technology* 34 (4): 1607–22. https://doi.org/10.1007/s13347-021-00477-0.

World Justice Project. 2023. *Rule of Law Index 2023*. Washington, DC: World Justice Project. https://worldjusticeproject.org/rule-of-law-index/downloads/WJPIndex2023.pdf.

Zuboff, Shoshana. 2019. "Surveillance Capitalism and the Challenge of Collective Action." *New Labor Forum* 28 (1): 10–29. https://doi.org/10.1177/1095796018819461.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches évaluées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

Research Director, Transformative Technologies Tracey Forrest
Director, Program Management Dianna English
Program Manager Erin Chreptyk
Publications Editor Lynn Schellenberg
Graphic Designer Abhilasha Dewan

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
www.cigionline.org