GEOPOLITICS OF AI STANDARDISATION AND ITS IMPACT ON FUNDAMENTAL RIGHTS

By

Gyan Prakash Tripathi

Submitted to

Central European University, Vienna

Department of Public Policy

In partial fulfilment of the requirements for the degree of

Master of Arts in Public Policy

In supervision of

Cameran Hooshang Ashraf, Ph.D.

Vienna, Austria

2025

COPYRIGHT NOTICE

[1]

For bibliographic and reference purposes this thesis should be referred to as:

Tripathi, Gyan Prakash. 2025 Geopolitics of AI Standardisation and its Impact on Fundamental Rights. MA thesis, Department of Public Policy, Central European University, Vienna.

[1] Icon by Font Awesome.

AUTHOR'S DECLARATION

I, the undersigned, **Gyan Prakash Tripathi**, candidate for the MA degree in Public Policy declare herewith that the present thesis titled "Geopolitics of AI Standardisation and its Impact on Fundamental Rights" is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography.
I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright.

I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 03 June 2025

Gyan Prakash Tripathi

ABSTRACT

Artificial Intelligence (AI) has become a transformative and disruptive force that reshapes consumer behaviour, economic paradigms, power politics, and security logics in equal measure. Because algorithmic architectures, training data, and optimisation parameters embody the social and, therefore, political visions of those who design, develop and deploy them, AI is far from neutral. It confers decisive economic, military and ideological advantages, and states now treat the technical specification of AI systems as a strategic domain. Echoing earlier struggles over Internet protocols, governments deploy international standard-setting bodies as instruments of geopolitical influence, hoping to entrench their normative preferences in the very "bare bones" of AI.

This thesis examines how the United States, the European Union, and China - alongside large middle powers viz., India, Australia, and South Korea - contest, negotiate, and diffuse AI standards within standard-setting organisations. I have synthesised realist, liberal-institutionalist and constructivist insights into a competitive norm-diffusion model, to reveal a sovereignty–interoperability paradox where states seek digital sovereignty even as global interoperability demands shared baselines. Democratic coalitions have embedded fundamental rights language in most new standards, yet authoritarian-leaning proposals persist in surveillance-heavy domains, sowing the seeds of a potential dual-centric bifurcation. Therefore, I conclude that AI standardisation now functions as a non-conventional but potent vector of geopolitical power.

Keywords: artificial intelligence, standardisation, norm-diffusion, fundamental rights, geopolitics.

for my *Dost*

ACKNOWLEDGEMENTS

I am deeply grateful to my supervisor, Cameran Hooshang Ashraf, Ph.D., for his guidance and motivation. I would also like to express my gratitude to the academic and non-academic staff at the Department of Public Policy, Central European University, Vienna.

I am very thankful for my friends, both old and new, who provided distractions when needed and encouragement when it seemed impossible to continue.

Lastly, I would like to express my deepest affection and indebtedness to my *Dost* Shri B K Tripathi, my parents Arti Tripathi and Prof. (Dr.) V K Tripathi, and my younger brother Divy Tripathi for their unwaivering support and unconditional love.

TABLE OF CONTENTS

INTRODUCTION

Artificial intelligence is increasingly woven into the fabric of everyday life, powering everything from personalised online services to critical infrastructure and defence systems. Because these systems rely on algorithmic models that reflect the data and design choices behind them, the rules that govern how AI is built and deployed—its technical standards—carry significant social and political weight. Standardisation seeks to create common vocabularies, performance benchmarks and assurance processes so that AI technologies can operate safely and interoperate globally (Cihon 2019). Yet the same standards shape how - or whether - fundamental rights, such as privacy, equality, non-discrimination, freedom of expression, and due process, are protected (Shaelou and Razmetaeva 2023). As AI becomes more pervasive, the way technical standards encode safeguards, accountability mechanisms and ethical principles will decisively influence whether this powerful technology ultimately advances or erodes human dignity.

As AI can provide crucial economic, military, and political advantages, its integration into society and politics will be reinforced and even amplify prevailing power systems (Horowitz 2018). Therefore, we see nation-states trying to use AI standardisation as a tool to exert their geopolitical influence and maintain (or establish) their supremacy in the AI global race. As the learnings from the not-so-long-ago past of Internet governance have shown, standards that provide the specifications for critical digital infrastructure administer, control, and ultimately govern the Internet (DeNardis 2009; Cameran Hooshang 2015). AI is not neutral or agnostic. The AI bare bones- the algorithms, dataset and training parameters decide, and ultimately mirror the social and hence political visions of nations that control them (Crawford

2021). As a result, it is a given that the ongoing "AI race" will necessarily dominate and impact the socio-political order for decades to come.

Technical standardisation helps to operationalise features such as interoperability, network compatibility and innovation (Busch 2011; Blind 2009). As (von Ingersleben-Seip 2023) points out, given the relatively green field in AI standards setting, while experts from geopolitical rivals are willing to cooperate in transnational standard-setting organisations, governments are not ready to collaborate on global ethical AI standards within international organisations.

RESEARCH DESIGN AND METHODOLOGY

Case Selection

Artificial intelligence (AI) standardisation has emerged as a new frontier of geopolitical competition, with profound implications for fundamental rights. Over the past five years, the United States, the European Union, and China – alongside large middle-power democracies such as India, Australia, and South Korea – have increasingly sought to shape global AI standards in line with their strategic interests and values.

The choice of the US, EU, and China as focal actors is justified by their outsized influence: together, they account for the largest AI developers, markets, and standardisation contributions. Their contrasting political systems and policy approaches present a rich comparative terrain for how fundamental rights might be championed or downplayed. Including India, Australia, and South Korea provides a necessary corrective to a purely great-power narrative. These three represent different regions (South Asia, Oceania, East Asia). Each has shown normative agency: India as a leader of the Global South with democratic credentials, Australia as a close ally of the US with an independent streak in regional diplomacy, and South Korea as a technologically advanced democracy often navigating between the US and China. All three are members of relevant forums – e.g. all are in SC 42 (participating members), in GPAI (founding or early members) (PIB 2022), and involved in multi-country partnerships (India and Australia in Quad, South Korea in OECD and a participant in G7 outreach). Their inclusion enables the thesis to explore whether and how *norm convergence* can be extended beyond the West, and how these states participate in or choose sides in standards contests.

The *research puzzle* driving this thesis is how these powers contest, negotiate, and diffuse AI standards across multiple venues, and what consequences this has for fundamental rights, including privacy, equality, non-discrimination, assembly and association (Cameran 2020), due process, and religion (Cameran 2022). This puzzle is significant not only because technical standards can "embed [countries'] values" into technology, but also because the way AI systems are governed internationally will shape whether human rights are protected or eroded in the algorithmic age.

Research Questions

The following empirical research questions have guided this thesis:

1. How have the US, EU, and China each advanced their preferred AI governance norms through international standard-setting bodies from 2020 to 2025?
2. What roles do large middle-power states such as India, Australia, and South Korea play in bridging or amplifying the normative contest between the major powers?
3. To what extent are fundamental rights principles reflected in the emerging AI standards, and how do geopolitical dynamics influence these outcomes?

While examining these questions, I also considered whether current trends indicate a shift toward polycentric convergence on AI standards or a dual-centric/bipolar fragmentation of standards aligned with competing geopolitical blocs. Through this, I aim to highlight the power politics of AI standards and their impact on rights, at the intersection of international relations, technology governance, and human rights scholarship.

Methodology

This research employs a qualitative case study methodology to analyse data from publicly available documents and examine the geopolitics of AI standardisation. The approach is qualitatively driven, making it suitable for unpacking complex socio-technical processes and normative content that quantitative measures cannot capture.

As I have elaborated above, the four case studies were chosen through a purposive, theory-driven sampling to illustrate different facets of the competitive norm diffusion model. By covering major powers and large middle powers, as well as a range of institutional venues (technical standard bodies, multilateral bodies, multi-stakeholder initiatives, and diplomatic forums), the cases capture the polycentric nature of AI governance, as well as instances of both convergence and contestation on rights issues.

The primary mode of data collection was document analysis. I examined a corpus of documents including standards drafts and final texts, official strategy papers and policy documents (EU Commission's AI White Paper 2020; China's New Generation AI Development Plan and AI governance white papers; U.S. Department of Commerce/NIST framework documents; G7/G20 ministerial declarations on digital policy), and plenary meeting minutes or press releases from relevant bodies.

Given that many standards negotiations occur behind closed doors or in technical jargon, contextual documents and expert analyses were also consulted. For example, to interpret the significance of an ISO voting result or an ITU proposal, I relied on contemporaneous reporting from trusted outlets and policy briefs by experts that provided authoritative insights

into these matters. These secondary sources helped infer the positions of actors in instances where official minutes are sparse.

Limitations and Reliability

In an ideal research scenario, I would conduct elite interviews with key stakeholders: EU Commission officials involved in standardisation, national delegates to SC 42 and ITU, industry representatives on IEEE standards committees, and civil society experts engaged in these forums. Such interviews would yield insights into motivations, negotiation tactics, and informal dynamics that documents cannot fully reveal. Further, technical documents and standards are either not public or require a heavy subscription fee. Therefore, due to access and timing constraints, I relied on material and scholarship that already existed.

Theoretical Framework

International efforts to govern AI via standards can be understood through a *"competitive norm-diffusion"* model, which integrates realist, liberal-institutionalist, and constructivist perspectives. In this framework, states are viewed as both power-seeking actors and norm entrepreneurs, operating within institutional constraints while attempting to disseminate their preferred norms globally. This section outlines the theoretical underpinnings of the model and defines key concepts used in my analysis, notably *polycentric vs. duo-centric governance* and a typology of norm diffusion strategies (offensive, defensive, bridging).

Realist Lens: Power and Sovereignty

A realist view emphasises that major states engage in AI standards-setting to pursue relative gains, national security advantages, and technological sovereignty. Control over

standards is seen as a form of power: by setting the rules, a state can *"potentially embed that country's values"* and advantages in the global technological infrastructure (Schmitt 2022).

From this angle, the US, China, and EU behave in strategic ways - China's heavy investment in standards leadership (sending large delegations, placing officials in key committee roles) is interpreted as a bid for hegemony in emerging tech, consistent with its broader ambition to become an "AI superpower" (Matthews 2025). The US historically relied on its market dominance to promulgate de facto standards (the dominance of English-language coding standards, IEEE protocols, etc., often mirrored US values like open Internet) (Techatassanasoontorn and Suo 2011). The EU, though not a traditional military power, exercises what might be termed *"normative power Europe"*, using its regulatory clout as leverage – realism adapted to the regulatory state (Ekdal 2021). Realism also highlights the concept of duo-centric or bipolar structures: if two superpowers (e.g. the US and China) dominate spheres of influence, one could see the formation of two separate blocs of standards.

Liberal-Institutionalist Lens: Institutions and Interdependence

Liberal institutionalism, by contrast, notes that repeated interactions in international organisations and the pursuit of shared economic benefits (or absolute gains) from common standards can mitigate pure power politics. Standards bodies like ISO, IEC, ITU, and IEEE are longstanding institutions with established rules (including consensus procedures, voting rules, and multi-stakeholder participation in some cases) that shape the behaviour of actors. Even powerful states must build coalitions and adhere to procedural norms to have standards approved. This lens explains instances of cooperation, such as the US and EU collaborating in the Trade & Technology Council to harmonise AI terminology and risk management approaches, suggesting that institutional dialogue can produce convergence despite competitive

impulses. Liberal theory also underpins the idea of polycentric governance. A *polycentric* governance system is one with multiple centres of decision-making authority that are autonomous yet interdependent (Ostrom 2010). In AI standardisation, polycentricity is reflected in the diverse network of institutions – each handling different aspects (e.g., ISO handling technical specifications, OECD handling high-level principles, UNESCO handling ethical guidelines, and the G20 providing political endorsement).

Constructivist Lens: Norms and Identity

Constructivism focuses on how state identities and beliefs shape their preferences, and how norms diffuse through socialisation. In this view, the content of the standards, e.g. whether privacy is protected, whether "trustworthiness" is defined to include human rights, is a product of ideational influence as much as material power. The EU's self-identity as a *"regulatory superpower"* and guardian of human rights leads it to promote norms like human-centric AI and "trustworthy AI" in every forum (Collins et al. 2023). China's political culture and governance model emphasise state control, social order, and collective security; these normative priorities manifest in the type of standards it proposes (e.g. standards enabling surveillance and centralised authority over networks) (Comerma 2024).

Importantly, constructivism introduces the mechanism of norm diffusion. Norm diffusion can be *direct* (powerful states persuading or coercing others to adopt their standards) or *indirect* (norms spreading through emulation, professional networks, or market forces) (Villarino 2023). The "Brussels Effect" is an example of indirect diffusion through market mechanisms and legal emulation (Bradford 2020). By contrast, China lobbying developing countries in ITU to vote for its proposals, sometimes bundled with Belt-and-Road tech deals, is

a more direct diffusion attempt via inducements and socialisation into China-led initiatives (Heeks et al. 2024).

Polycentric vs. Duo-Centric Governance

Within this model, *polycentric governance* refers to a structure in which multiple independent centres of authority coexist and collaborate to varying extents. In AI standardisation, this means a world where ISO, ITU, IEEE, OECD, and other organisations all function as nodes contributing to a tapestry of global AI norms. Polycentric governance is often associated with convergence tendencies – because if many forums are working on related issues, there is an opportunity for overlap and common reference points (for example, ISO's standards might refer to OECD definitions, IEEE might adopt ISO terminology, etc., as is already happening) (Angst et al. 2022). *Duo-centric governance*, on the other hand, implies a bifurcated structure with two predominant poles that have their own relatively self-contained ecosystems. A duo-centric outcome in AI might look like this: one pole (e.g. US/EU and allies) adheres to standards that require vigorous fairness testing, transparency, human oversight (mirroring EU–OECD principles), while another pole (China and partners) uses standards that emphasise security, efficiency, and government prerogative, with minimal privacy constraints (Feakin 2025). Interoperability between these systems could wane, as devices and algorithms are built to different specifications or with different ethical guardrails. Data flows could also be restricted along bloc lines due to incompatible regulations (for instance, Western countries restricting AI software exports to nations that don't meet specific human rights safeguards, or vice versa).

CHAPTER I: CURRENT STATE-OF-THE-ART

The growing body of literature on AI global governance highlights a fundamental tension between the rapid deployment of AI technologies and the underdevelopment of international norms to guide them (Schmitt 2022). Early analyses catalogued the proliferation of AI ethics guidelines and principles in the late 2010s, noting convergence around ideals such as transparency, fairness, accountability, and human-centricity (many inspired by the International Bill of Human Rights), but divergence in implementation and enforcement. More recent governance analyses emphasise the shift from principles to practice. For example, the EU's AI Act has been examined as a pioneering effort to translate ethical principles into law, with commentators assessing its risk-based approach and potential global impact. At the same time, strategists in the US have argued that AI competition is fundamentally *"a values competition"* (McInerney 2024), suggesting that which country or model leads in AI will determine whether liberal-democratic or authoritarian values prevail in digital norms. The U.S. National Security Commission on AI's 2021 Final Report explicitly framed the AI rivalry with China in ideological terms: "We must…build privacy-protecting standards into AI technologies and advance democratic norms to guide AI" (National Security Commission on Artificial 2021). This perspective aligns with the concept of *"technological competition as normative competition,"* indicating that AI governance is not value-neutral but deeply political (McGeachy 2019).

A rich literature exists in international political economy and international relations on how standards-setting can confer a competitive advantage and reflect state power (Mattli 2001; Egan 2002; Mattli and Büthe 2003; Austin 2000). Classical works note that whoever sets technical standards can gain economic leverage ("winning" standards can lock in markets for

national industries) and also diffuse their regulatory philosophies (Tassey 2000). Recent analyses focus on China's concerted push in this domain. Beijing's "Standards Power" doctrine and the anticipated *China Standards 2035* plan underscore an official strategy to dominate emerging tech standards, including AI (Ford 2022). As (Olson 2020b) observes, Chinese experts frequently occupy leadership positions in standards bodies, and Chinese companies often send the largest delegations, seeing standards-setting as *"seizing the commanding heights"* of tech innovation (Kania and and Costello 2021). Conversely, Western governments have become alert to the geopolitical stakes of standards only recently – for instance, the G7 Digital and Technology Ministerial Declaration 2021 created a *Framework for Collaboration on Digital Technical Standards* (G7 2021), warning that standards for critical tech could affect "*shared values as open and democratic societies*".

Two key concepts in the literature are forum shopping and standard fragmentation. Forum shopping refers to actors choosing favourable venues to advance standards or rules they prefer. For example, some scholars note that China has favoured the ITU for pushing specific Internet and surveillance standards because the ITU operates on a one-country-one-vote intergovernmental model (where Beijing can leverage developing country support) and lacks the multi-stakeholder checks present in bodies like the Internet Engineering Task Force (IETF) or IEEE (Sherman 2022). Meanwhile, the US and EU often prefer private-sector-led or consensus-based bodies (IEEE, ISO, OECD) where democratic norms are embedded in processes (The White House 2021). This fragmentation of forums leads to what scholars describe as a *"regime complex"* for AI governance – a decentralised, polycentric landscape with overlapping initiatives and no single authority. (Schmitt 2022) maps this nascent regime and finds it *"polycentric and fragmented"*, yet observes early signs of consolidation around

fora like the OECD which provide epistemic authority. Indeed, the OECD's 2019 AI Principles, which were the first intergovernmental AI standards – have been influential beyond the OECD, gaining G20 endorsement and inspiring many national frameworks (OECD 2019). This suggests a potential *convergence node* in the regime complex, even as competition persists elsewhere.

A relevant strand of literature concerns the extraterritorial impact of the EU's regulations – termed the *Brussels Effect* (Bradford 2020). In the context of AI, researchers ask whether the EU's AI Act could globalise EU-style norms, as GDPR did for privacy. The notion of a *"Beijing Effect"* is also emerging, denoting China's ability to export its digital governance norms (e.g. through the Digital Silk Road initiative) (Mueller and Yoo 2023). Empirical studies have noted how Chinese surveillance technologies and accompanying standards (for facial recognition, city CCTV systems, etc.) are adopted in parts of Africa, Asia and Latin America (Jili 2022; Warner and Ajibade 2024; Melson 2021). This export is often accompanied by Chinese-led training and standardisation assistance, embedding technical norms that may lack privacy safeguards. The literature thus identifies *duelling diffusion processes*: a liberal diffusion of human-rights-oriented standards versus an authoritarian diffusion of surveillance-friendly standards, raising the prospect of a *standards schism* or "splinternet" in AI if each gains traction in different regions (Gabbott 2023; Bazoobandi et al. 2025).

Scholars and advocates have increasingly applied human rights law as a lens for evaluating AI systems. Articles in legal scholarship highlight that AI can infringe upon privacy rights (through mass data processing and surveillance), equality (via biased algorithms that discriminate), freedom of expression (through content moderation or generative AI that shapes discourse), and due process (when automated decisions lack transparency or recourse). The

"International Bill of Human Rights" (UDHR, ICCPR, ICESCR) provides a normative baseline that some argue should be explicitly incorporated into AI governance. For instance, Access Now and Human Rights Watch have called for a moratorium on specific AI applications (such as autonomous facial recognition) that are incompatible with the right to privacy and freedom of assembly. Academic works (e.g. by (Latonero 2018; Prabhakaran et al. 2022) promote a *"human rights by design"* approach in AI development, aligning with the idea of *embedding rights in technical standards*. UNESCO's 2021 Recommendation on AI Ethics is a milestone in formalising this approach at a global scale – it explicitly references human dignity, human rights and fundamental freedoms as cornerstones, and it was adopted unanimously (UNESCO 2022). Notably, even China and Russia – often criticised for domestic human rights abuses – joined consensus on that Recommendation, though their commitment to its implementation remains questionable. This points to a recurring theme in the literature: a gap between high-level normative agreement and on-the-ground technical standards or practices. While human rights language may be agreed upon in principle (often due to diplomatic pressure and reputational incentives), translating those principles into concrete standards and certifications is fraught. Analysts highlight that industry-led standards often overlook human rights considerations unless there is strong advocacy from civil society or the government.

Based on this review, two gaps stand out. First, there is limited scholarly examination of international standards bodies (ISO, ITU, IEEE) as contested sites of AI governance. While policy reports have noted Chinese proposals at the ITU or US/EU cooperation in ISO, academic analysis of the negotiations, power plays, and normative content in these bodies is scarce. This thesis contributes by illuminating how standards politics unfolds and how forum shopping operates in practice. Second, the nexus between geopolitics and fundamental rights in AI

remains underexplored. Existing work often treats ethical AI principles in general terms or focuses on the law of one jurisdiction. Still, few have analysed how geopolitical rivalries might shape the global *substance* of AI norms in relation to rights. By interrogating multiple powers' approaches across several forums, this research sheds light on whether global AI standards are trending toward reinforcing human rights or diluting them under geopolitical pressures.

In doing so, it builds on and synthesises insights from international relations theory, technology policy, and human rights law. I have attempted to operationalise theoretical ideas (like norm diffusion and regime complexes) in the concrete domain of AI standards. This interdisciplinary approach is relatively novel in the literature, which is often siloed into either high-level ethical discussions, technical standards discourse, or International Relations power analysis. By bringing these together, I provide a more comprehensive understanding of the geopolitics of AI standardisation and to inform both scholarship and policy on how to navigate the delicate balance between technological sovereignty and global interoperability in upholding fundamental rights.

CHAPTER II: BRUSSELS EFFECT 2.0 - EXPORTING TRUSTWORTHY AI NORMS

The European Union has leveraged its regulatory power and normative agenda to influence international AI standards and norms, a phenomenon dubbed as *"Brussels Effect 2.0"* (Engel and Groussot 2025). It traces the EU's promotion of a "trustworthy AI" framework from 2018 onwards and its impact on two key targets: the ISO/IEC JTC 1 SC 42 standards (notably the AI management systems standard, ISO/IEC 42001) (Directorate-General for Communication 2024), and China's evolving AI governance discourse. Realist motives (maintaining regulatory primacy) dovetail with constructivist identity (the EU as a values leader) to drive this strategy, executed through institutional routes (ISO committees, multilateral principles) consistent with liberal institutionalism. The result is partial convergence around EU-inspired norms, illustrating the action of offensive norm diffusion.

EU's Trustworthy AI Framework

The EU's normative push began in earnest with the High-Level Expert Group on AI, which in 2019 released *Ethics Guidelines for Trustworthy AI* (High-Level Expert Group on AI 2019). These guidelines defined *"trustworthy AI"* as AI that is lawful, ethical, and robust, and articulated seven key requirements (including privacy and data governance, non-discrimination, transparency, human agency and oversight, accountability, etc.). This effectively operationalised fundamental rights principles into an AI context. The *European Commission's White Paper on AI (February 2020)* built on these ethics guidelines and floated the idea of regulating "high-risk AI" with requirements for trustworthiness (EU Commission 2020). Commission President Ursula von der Leyen explicitly invoked creating a "GDPR for AI" (Feldstein 2024) – signalling the intent to export EU norms via market power. The *"Brussels*

Effect" in AI, as analysed by researchers, posits that EU regulation will have an extra-territorial impact because companies globally will adjust products to meet EU standards rather than forgo the EU market (Bradford 2020).

Even before legislation, the EU set about internationalising its AI vision. It spearheaded the drafting of the OECD AI Principles (2019), ensuring they mirrored EU values (indeed, principles like *"AI should respect the rule of law, human rights and democratic values"* are clearly aligned to Europe's approach). When the G20 quickly endorsed these OECD principles in mid-2019, it marked the first global political acceptance of the *"trustworthy AI"* paradigm, with an emphasis on people-centric and rights-respecting AI (OECD 2021). Notably, China was part of the G20 consensus. Chinese officials and scholars took note: A Chinese state think tank's *White Paper on Trustworthy AI* (published by CAICT in 2021) opens by stating, *"The development of trustworthy AI is becoming a global consensus"*, citing the G20 Principles and noting that *"the EU and United States have placed the enhancement of user trust and development of trustworthy AI at the core of their AI ethics and governance"* (China Academy of Information and Communications Technology 2021). This is a striking example of normative diffusion – Chinese experts explicitly acknowledging and internalising the EU/U.S. framing of AI governance.

China initially approached AI governance from a tech-development standpoint (with its 2017 New Generation AI Plan focusing on innovation and only cursory mentions of ethics). However, as global norms coalesced around "trustworthy AI," China adapted to avoid isolation. The 2019 Beijing AI Principles (proposed by a Chinese research consortium) and subsequent government-affiliated guidelines in 2019–2021 started using language like *"AI for Good"*, *"fairness"*, and *"privacy protection"* ('Beijing Artificial Intelligence Principles' 2019). The

July 2021 white paper by the China Academy of Information and Communications Technology is a concrete example of China internalising global AI governance norms (China Academy of Information and Communications Technology 2021). It frames trustworthy AI as essential to resolve a "crisis of trust" in AI and explicitly references the G20/OECD principles as well as EU/U.S. efforts. The white paper endorses concepts like *privacy-preserving machine learning (federated learning, differential privacy)*, *algorithmic transparency*, and *avoidance of bias*, which are cornerstones of the Western-led AI ethics discourse. It even calls for *drafting Chinese legislation on trustworthy AI* and developing *AI audit mechanisms*, steps clearly paralleling EU regulatory thinking.

This represents *norm diffusion through socialisation and competitive emulation*: Chinese actors saw the writing on the wall that trust and ethics were becoming prerequisites for AI's global acceptance and decided to not reject these norms but rather claim a stake in them. Of course, China's interpretation often emphasizes *"controllability"* and *"security"* alongside fairness – the CAICT paper lists *"controllable, reliable, transparent, and explainable AI with privacy protection, clear responsibilities, and diversity and tolerance"* as goals (China Academy of Information and Communications Technology 2021). The inclusion of "controllable" reflects the CCP's focus on ensuring AI systems do not threaten social stability or Party oversight. Nevertheless, it is significant that *privacy protection* and *fairness* are explicitly mentioned – concepts rooted in human rights discourse that historically were muted in Chinese policy. The white paper acknowledges that *"all walks of life around the world attach great importance to trustworthy AI"* and that turning principles into practice is an *"urgent task"* for industry. This acknowledgement validates the success of EU (and allied) norm

entrepreneurship: it shaped what is considered the global consensus to which even China feels compelled to pay lip service and partially adhere.

The Brussels Effect 2.0 in Action: Scholars have termed this "2.0" because it's not just de facto market influence (as with GDPR) or an extension of it, but a conscious strategy by the EU to work through international bodies to diffuse its norms. All these moves constitute a deliberate normative diffusion campaign. The impact is clearly seen: terms like "trustworthy AI", "human-centric AI", and "ethics by design" – once emanating from European policy circles – are now part of the lexicon in global standards discussions, company marketing materials worldwide, and even Chinese policy documents. The competitive aspect is that EU's push in some ways forced others to react. U.S. tech companies, wary of hard EU law, supported voluntary standards with similar ethos (to pre-empt regulation), which aligns with EU's end goals albeit via different means. China, concerned about losing legitimacy, incorporated much of the language into its own standards narrative, as discussed. This dynamic shows the power of an offensive norm diffusion strategy when backed by significant market and regulatory weight.

From a theoretical stance, this case underlines constructivist insights (norm entrepreneurship and socialisation) and liberal institutionalism (the use of forums like ISO and OECD to lock-in norms). Realism is not absent either: the EU is securing a leadership role that ensures its industries face less adjustment when global standards align with EU rules, and it pre-empts rivals from imposing a conflicting vision. We now turn to the second case, where I examine transatlantic dynamics – effectively the partnership and sometimes friction between the US and EU as they navigate converging around those norms in practice, which will shed

light on how solid the “trustworthy AI” consensus truly is when operationalising it across different regulatory philosophies.

CHAPTER III: US–EU - TRANSATLANTIC ALIGNMENT AND COMPLIANCE CULTURES

In this chapter, I explore the interplay between the United States and the European Union in shaping AI governance through a risk-based approach, highlighting both convergence on high-level principles and divergence in regulatory and compliance cultures. It contrasts two landmark initiatives – the U.S. *NIST AI Risk Management Framework (RMF)* and the EU's *AI Act* – as exemplars of each side's approach. It examines how transatlantic cooperation (particularly via the Trade and Technology Council, TTC) has led to a shared vocabulary of *"trustworthy AI"* and *risk management*, even as differences remain in enforcement and scope. The case illustrates a combination of norm diffusion (the EU's influence on US thinking, and vice versa in certain technical areas) and negotiation to reconcile distinct legal and political traditions. It also illustrates how two liberal actors navigate power dynamics differently from the adversarial US–China or EU–China context: here, the challenge is aligning systems to avoid contradictory standards that could hamper innovation and trade, all while upholding their shared values of fundamental rights.

By 2020, both the US and EU had publicly committed to *"trustworthy AI"* as a goal, thanks in part to the OECD/G20 Principles they co-sponsored (G20 Digital Economy Ministers Meeting 2020). Both recognised that not all AI carries equal risk and that governance should be proportionate to potential harm. The EU formalised this in the *AI Act* by categorising AI uses into risk levels – *unacceptable risk* (banned practices like social scoring and real-time biometric surveillance in public), *high-risk* (like AI in recruitment, lending, law enforcement, which must meet strict requirements), *limited risk* (like chatbots or deepfakes, requiring transparency), and *minimal risk* (little to no regulation for most AI) (*Regulation (EU) 2024/1689* 2024). This

schema itself was influenced by earlier U.S. thinking on risk tiering in areas such as medical devices, as well as by industry input. Conversely, the US did not at first have a single comparable instrument. Still, various agencies produced guidelines (e.g. FAA on autonomous drones, FDA on AI in medical devices) and the White House OSTP in late 2020 issued a policy memo with *10 principles for federal use of AI* (e.g. fairness, transparency) – these were non-binding but signaled acknowledgement of similar issues as the EU's (Hao 2020).

The NIST AI Risk Management Framework project, launched in 2021 and culminating in version 1.0 in January 2023 (Tabassi 2023), became the focal point of U.S. efforts to articulate a cohesive approach. From the outset, NIST coordinated with European counterparts. Indeed, by late 2022, a *TTC Joint Roadmap on Evaluation and Measurement for Trustworthy AI* explicitly aimed to align NIST's work with the EU's regulatory developments ('TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management' 2022). A key statement in that roadmap: *"The United States and EU acknowledge that a risk-based approach and a focus on trustworthy AI systems can provide people with confidence in AI... Both parties are pursuing risk-based approaches that operationalize these values."*. It also references shared dedication to the OECD AI Recommendation and democratic values, grounding their collaboration in normative common ground.

Consequently, NIST's AI RMF and the EU AI Act ended up sharing many foundational concepts and even terminologies (thanks to a *glossary alignment* effort in the TTC). Both, for example, emphasise *socio-technical factors*, bias mitigation, transparency, robustness, and human oversight in high-risk contexts. Both frameworks define *trustworthiness characteristics* such as safety, explainability, fairness, privacy, and security. This could have been penned in Brussels as easily as in Washington. We see here a clear case of *constructivist norm alignment*

— the US integrating human rights and ethics language into what it frames as a voluntary standard. Europe's imprint is undeniable; in earlier years, U.S. agencies spoke less of "rights" and more of "innovation" and "AI leadership". By 2022, even the U.S. *Blueprint for an AI Bill of Rights* (an OSTP document published October 2022) (OSTP 2022) speaks of ensuring AI *"protects civil rights, civil liberties, privacy"*, covering issues like algorithmic discrimination, data privacy, and notice and explanation for AI-driven decisions. This Blueprint is non-binding but politically significant, and was cited alongside NIST RMF in the TTC Roadmap as a complementary U.S. example.

Despite high-level convergence, the US and EU approached compliance in distinct ways, rooted in their regulatory philosophies and political systems. The EU AI Act is *hard law*, it will impose legally binding requirements on providers and users of high-risk AI systems. These include conducting *conformity assessments*, documenting compliance (technical documentation, logging, etc.), and potentially involving third-party audits (notified bodies) for some AI systems. Non-compliance could lead to hefty fines (up to 6% of global turnover, akin to GDPR fines) – an enforcement mechanism external to the company. The Act also creates prospective institutions (like an EU AI Board, national supervisory authorities) to oversee implementation, reflecting the EU's preference for formal regulatory oversight. This embodies a compliance culture of ex ante rules and external accountability.

In contrast, the NIST RMF is *voluntary guidance*. It provides a structured process: *"Map, Measure, Manage, Govern"* functions that organizations can adopt to self-regulate AI risks (Ricciardi Celsi and Zomaya 2025). It explicitly states it is *"non-sector-specific, use-case agnostic, and voluntary"*, meant to be adaptable. Compliance in the U.S. approach is largely market-driven or internally driven - the idea is that companies will implement the RMF if they

want to earn trust from consumers or avoid future liability, and that government agencies might require their vendors to use it (thus incentivising adoption) (Kulothungan, Mohan, and Gupta 2025). There are no fines directly tied to the RMF, and no AI-specific regulator at the federal level. Instead, the U.S. relies on existing laws (e.g. non-discrimination law, consumer protection by FTC, etc.) and general oversight to catch egregious harms. This reflects the American *"light-touch"* regulatory tradition and trust in industry-led standards.

This difference in compliance culture – *mandatory vs. voluntary, external enforcement vs. self-governance* – leads to practical divergence (Engler 2023). One example is how bias mitigation is treated. The EU AI Act will require testing for bias before AI is deployed in high-risk areas and mandates representative datasets "to the extent possible" to minimize discrimination, with documentation to be available to authorities (*Regulation (EU) 2024/1689* 2024). Under NIST RMF, bias mitigation is strongly recommended (it defines "harmful bias" and suggests impact assessments), but if a company merely partially implements it or does so internally without disclosure, there's no immediate penalty – unless an incident triggers FTC or EEOC action under existing anti-bias laws. As a result, a software developer deploying an AI system in both EU and US markets might face stricter pre-deployment steps in the EU (or with EU customers) than in the US. (Engler 2023) also emphasises that *"the specifics… have more differences than similarities"*, pointing out especially in socio-technical domains like hiring or online speech, the EU is legislating obligations whereas the US is not. It warns of *"significant misalignment"* in such sectors if current trajectories hold.

In terms of the competitive norm diffusion typology, the EU was initially *offensive*, pushing its comprehensive regulation. The US response was partly *defensive* (resisting any notion that it should copy EU law) but also *bridging*, working to find common principles to

ensure a united democratic front. The TTC's very existence is a liberal institutional mechanism to avoid a standards war between allies that would only benefit authoritarian competitors. Both sides realised that a transatlantic split in AI governance would weaken the West's position in setting global norms, potentially allowing China to exploit differences or sell its AI claiming "the West can't even agree, so follow our simpler model." Thus, strategic considerations pushed the US and EU toward convergence – a liberal self-coordination to bolster normative power collectively. We saw this as they presented a united front in international discussions (e.g., joint proposals in OECD about AI classification, or co-leadership in GPAI working groups).

Realism is present in that the US and EU each want to ensure their industries flourish. The US was keen to ensure EU regulation doesn't unfairly lock out US companies or become a trade barrier. Part of the TTC's mission is to ensure regulatory approaches are at least *equivalent* or mutually recognized to avoid hindering transatlantic trade. For example, will the EU accept an AI system tested under NIST methods as meeting EU standards? If yes, US companies avoid duplicate compliance processes. The joint roadmap hints at pursuing that kind of equivalence ('European Union: Initiatives on AI and Emerging Technologies US-EU TTC' 2022). From the EU perspective, having the US adopt similar standards voluntarily would vindicate its normative leadership and ensure European companies operating in the US aren't undercut by lower standards. So a balance of power and interest led to cooperation – a case of realist interests (maintaining an edge over China) aligning with institutionalist win-win (ease trade) and constructivist alignment (shared values).

CHAPTER IV: CHINA AT THE ITU - THE FACIAL RECOGNITION STANDARDS BATTLE

This chapter delves into a high-profile contest within the International Telecommunication Union (ITU) – specifically the *ITU-T Study Group 16* proposals on facial recognition and surveillance standards spearheaded by Chinese entities between 2019 and 2022 (Teleanu 2021). Here, I analyse the concerted push towards new international standards for facial recognition by Chinese tech companies (backed by the Chinese government), and Western governments, along with human rights advocates who pushed back due to concerns over privacy, surveillance, and potential abuse (such as ethnic profiling) (Wouters 2023). This case exemplifies and aligned countries advocating a technically oriented but values-laden standard, versus a coalition of liberal democracies resisting the encroachment of authoritarian norms into global standards. It highlights forum shopping (China leveraging the ITU's state-driven structure rather than multi-stakeholder venues) and illustrates the risks that divergent values pose to global interoperability of standards. Fundamentally, it shows how standards that seem purely "technical" can become proxy battlefields for human rights and governance models (Murgia, Yang, and Gross 2019).

Other scholars such as (Sukumar and and Basu 2024) have also highlighted how China and Russia have used the UN cybercrime negotiations at an Ad Hoc Committee (AHC) of the United Nations through proposals pushing "for states to have greater control over digital infrastructure and activities". The ITU's Telecommunication Standardization Sector (ITU-T) is one of the oldest international standard bodies, traditionally focused on telecom protocols and infrastructure. Unlike ISO or IEEE, ITU is a UN agency and every member state gets a vote; sector members (companies) participate but states have the final say in approving

recommendations. In the late 2010s, ITU-T broadened its purview to certain emerging tech standards including AI applications in telecommunications. During this time, Chinese influence at the ITU was at a peak, the Secretary-General of ITU (2015–2022) was Zhao Houlin of China, and Chinese companies were extremely active contributors (Lott 2022; Schaefer and Pletka 2022). China had declared its ambition to lead in global standards, and the ITU was a prime venue thanks to its state-centric nature.

In 2019, journalists obtained leaked ITU documents showing that Chinese companies like ZTE, Dahua, China Telecom, Huawei and others had submitted multiple proposals for new international standards on facial recognition, video monitoring, city and vehicle surveillance (Murgia, Yang, and Gross 2019). One proposal was a *"Face Recognition Application Profile"* that described methods for detection and classification of human faces in video systems – including attributes like ethnicity, skin color, and other personal features (Burt 2019). Another was a framework for *"smart street lamp surveillance systems"* integrating cameras, pitched by ZTE and China Mobile, which happened to align with those companies' commercial products (Lebryk 2021). The leaked specs revealed functions for identifying individuals or characteristics at distance, real-time processing, and database integration. In essence, these proposals aimed to standardise how surveillance cameras and facial recognition software interoperated globally, potentially opening markets for Chinese surveillance tech in developing countries by giving them an ITU "stamp of approval".

China's motivation was twofold: commercial (to create a global market where Chinese products set the benchmark) and ideological/strategic (to normalize high-tech surveillance as an acceptable practice, embedding its governance norms in global standards). As (Olson 2020b) notes, *" [Chinese companies] are active in helping to shape facial recognition and surveillance*

standards at the ITU. These standards are especially influential in determining how new technologies are deployed in developing countries.". Indeed, many developing countries look to ITU recommendations as guidance for national tech deployments. If the ITU blesses a certain architecture for facial recognition networks, a country in Africa or Asia might adopt it wholesale (and buy Chinese equipment to implement it).

When civil society learned of these proposals (initially through media exposés like the Financial Times in late 2019), alarm bells rang. They pointed out how the standards under discussion cross the line from mere technical specifications to policy recommendations that tilt towards an intrusive degree of surveillance, including potential ethnic and racial profiling. Western democracies, including the US, UK, Germany, and others, mobilised within the ITU to scrutinize and slow these proposals. Unlike in some other standards bodies, blocking a proposal in ITU can be done if enough countries raise concerns or if consensus isn't reached (ITU often works by consensus but can vote if needed). The EU countries and US were reportedly concerned that approving these without clear privacy safeguards would implicitly legitimate mass surveillance practices contrary to their values and possibly even their laws (like GDPR, which would conflict with systems that indiscriminately identify individuals in public). They likely filed comments asking for clarification on data protection, necessity, etc., and possibly insisted on "pause for study".

What is clear is that the *very attempt* by China had broader consequences: it galvanized Western governments to pay more attention to standards venues as arenas of strategic competition (the G7 statements on digital technical standards in 2021 were partly a reaction to episodes like this). It also showed the limits of global consensus: unlike the OECD principles (which China signed) or UNESCO AI ethics (which China also signed), when it came to

operationalising a specific technology, values clashed openly. This is supported by the analysis by (Sukumar and and Basu 2024) where they also conclude that "it is clear from Russia and China's proposals that both states view the proposed AHC convention on cybercrime as an opportunity to chart new rules for the road for cyberspace".

This ties to a broader Chinese strategy also seen in another ITU initiative, *"New IP"*, an attempt to create a new Internet protocol architecture more centralised and controllable by governments (Chen et al. 2020). Western governments vehemently opposed that and it was shelved (the ITU rejected "New IP" in 2020) (Haigh 2025; Yin 2024). The facial recognition standards saga is analogous: an authoritarian-flavored technical proposal meeting democratic resistance. It also pointed out the assumption that China's integration would make it more like the West was not holding; instead China is reshaping these institutions to fit its model.

From a fundamental rights perspective, had the Chinese proposals passed as originally written, it could have normalized the inclusion of features like *"recognition of ethnic features"* as a standard capability, implicitly legitimising such uses. It would not itself force countries to violate rights, but it creates a climate of acceptability and ease for those who want to (for example, an abusive regime could say they are just following international standards when implementing pervasive face-scanning of minority populations). Conversely, the blocking of these standards (or dilution) by the Western bloc can be seen as a defensive win for human rights norms – preventing an authoritative imprimatur on technology that undermines privacy and could enable discrimination (Ginsburg 2020).

It also typifies how China uses international bodies to *encode its norms into technical governance*. It had the foresight to recognise standardisation as a power tool, and this was an offensive norm diffusion attempt: pushing its governance model (ubiquitous surveillance,

population control tech) under the guise of technical specs. The opposing bloc used liberal institutional tactics (transparency, advocacy, alliance-building) to counter. Interestingly, unlike typical superpower showdowns, this played out mostly in diplomatic and technical arguments rather than public rhetoric; ITU processes are usually behind closed doors, so most public learned about it from leaks. This secrecy perhaps aided China initially, but once exposed, it became politically salient.

CHAPTER V: MIDDLE-POWER BROKERAGE IN AI NORM BRIDGING

In this chapter, we turn our focus on the roles of three influential large middle-power democracies viz., India, Australia, and South Korea, in the international AI governance landscape. Here I illustrate how these countries act as *"norm bridges"* and brokers, leveraging their unique positions to shape outcomes in forums such as ISO/IEC JTC 1 SC 42, the Global Partnership on AI (GPAI), and the Quad (Quadrilateral Security Dialogue) cooperation on technology. Each of these states has aligned broadly with the democratic, human-rights-oriented approach to AI, yet they also bring perspectives from the Asia-Pacific and Global South, aiding in building broader consensus beyond the US-Europe axis. They use their agency to forum shop and sometimes to mediate between great powers, contributing to a more *polycentric governance* and potentially mitigating the bipolar tendencies observed in prior cases. I examine specific contributions: India's leadership in multi-stakeholder AI initiatives and standardisation, Australia's proactive normative alignment and diplomacy, and South Korea's technical and diplomatic engagements, to show how these large middle powers punch above their weight in the diffusion and negotiation of AI norms.

India – "AI for All" and Strategic Bridging

India presents an interesting duality. It is the world's largest democracy, sharing values of freedom and rights in principle, yet it has its own sovereignty-driven approach to digital governance (insisting on data sovereignty, resisting outside interference) (Gaur 2025; Panday 2025; Derivry 2023). In AI, India's 2018 national strategy (*National Strategy for AI: #AIforAll*) set a tone of leveraging AI for inclusive growth and societal benefit, emphasizing principles like *"equality, AI for the benefit of all segments of society"* and identifying sectors like healthcare and agriculture for AI use (NITI Aayog 2018). While not a detailed ethical code, it

implicitly aligns with the idea that AI should not exacerbate inequality. India was a founding member of the Global Partnership on Artificial Intelligence (GPAI) in 2020, joining forces with the EU, US (which joined in 2020 after initially hesitating), and others to create a multi-stakeholder platform for responsible AI. By 2022, India took over as Chair of GPAI, indicating strong commitment (PTI 2022). Under India's chairmanship (2022–23), GPAI's focus included *AI for social good*, *responsible AI deployment in developing countries*, and bridging capacity gaps – aligning global AI ethics with development agendas (PIB 2022). This helped ensure that global AI discussions considered not just Western regulatory issues but also issues like *inclusivity, affordability, and cultural context*, thus broadening normative appeal and reducing the impression that "AI ethics" is a purely Western preoccupation.

In ISO/IEC JTC 1 SC 42, India is an active P-member through the Bureau of Indian Standards ('ISO/IEC JTC 1/SC 42 - Artificial Intelligence' 2017). Additionally, India's emphasis on *"AI for All"* resonated in UNESCO discussions: India strongly supported UNESCO's Recommendation on AI Ethics and pushed for mention of *equity and inclusiveness*, ensuring issues like digital divides were included as ethical considerations – something Western nations might not prioritise but developing nations do. By doing so, India acted as a broker between Western human-rights framings and developing world development-centric framings, helping achieve unanimous adoption.

Diplomatically, India leveraged its G20 Presidency in 2023 to highlight AI. Although 2023 was dominated by other issues, India's *"Delhi Declaration"* included notes on responsible AI and a plan to create a global digital public infrastructure repository – indirectly tied to AI governance and standards for public-good AI solutions (Vijayakumar 2024; G20 Delhi Declaration 2023). India's approach to facial recognition and surveillance is cautious:

domestically it deployed facial recognition for things like identifying missing children, but also faced criticism for police use in mass surveillance of protests (Sinha 2024). Yet, internationally, India did not back China's ITU standards – likely due to concerns over foreign dominance and its own civil society pressure. In ITU, India sometimes aligns with developing countries but broke from China/Russia on the "New IP" issue, siding more with the open Internet bloc in 2020. This suggests India doesn't automatically support Chinese tech norms if they clash with open society values or its independent stance.

Australia – Liberal Values and Regional Diplomacy

Australia has been a vocal proponent of democratic values in tech governance and often acts as a surrogate and partner to larger Western efforts, while also bridging into the Indo-Pacific region. It released an *AI Ethics Framework* in 2019 with 8 principles (including *human, social and environmental wellbeing; fairness; privacy protection; transparency; accountability*) (Dawson et al. 2019). These mirrored the OECD AI Principles (Australia was one of the 42 initial endorsers in 2019 and explicitly modeled its principles on OECD's). This alignment made Australia a credible advocate in the region for the OECD/GPAI normative approach. As a member of the OECD and GPAI, Australia actively participates in shaping guidelines and was among the first to integrate them domestically – thereby showing that adopting these standards is feasible, encouraging others in Asia-Pacific to do so.

Australia is also central to the Quad's technology collaboration. Within the Quad (Australia, India, Japan, US), Australia has championed initiatives on *critical and emerging tech standards*. At the Quad Leaders Summit (Sept 2021), it was announced that *"Quad Principles on Technology Design, Development, Governance, and Use"* were launched (Quad 2021b). These principles explicitly assert that tech should be shaped by *"shared values and*

respect for universal human rights". Australia's influence here is significant: as a large middle power, it lends credibility that these principles are not just a superpower (US) dictate but a collective stance including nations from the global West and East.

At ISO SC 42, Standards Australia (SA) is also P-member and has had experts in work on risk management and robustness standards. They also worked closely with allies (Australia being part of the "Five Eyes" intelligence alliance with the US, UK, Canada, NZ – those countries often coordinate in standards too to ensure security and rights are considered) (Kareem 2025).

Australia's normative stance is strongly pro-rights (its 2020 Human Rights and Technology final report even considered how to implement human rights impact assessments for AI domestically) (Australian Human Rights Commission 2021). Yet in practice Australia balanced this with security – e.g., Australia banned Huawei from 5G on security grounds and has urged neighbors to adopt secure network standards, showing it doesn't shy from excluding Chinese tech in critical areas to uphold trust and values of openness (IISS 2019). In global AI talks, this translates to pushing for standards that guarantee security (no hidden backdoors) and privacy.

South Korea – Technical Contributor and East-West Mediator

South Korea, a highly advanced tech economy, often plays a bridge between Western frameworks and Asian context. It is an OECD member (and contributed to the AI principles), and it joined GPAI at inception. South Korea also has its own AI ethics charter (the Korean Government announced "AI Ethical Standards" in 2020) (Park and Bang 2020), which includes principles like *human-centered, transparency, fairness, and safety*, very much echoing

OECD/EU values. However, South Korea is geographically and economically close to China, and thus must tread carefully. It typically aligns with democracies in values, but in forums it may take a softer approach to avoid antagonising China directly (Seoul tries to maintain stable relations with Beijing). This can make Seoul a useful mediator: it can talk to China in Asian cooperative language while still upholding democratic norms.

In ITU, for example, a South Korean, Chaesub Lee, was the Director of ITU-T Standardization Bureau (2015-2022) (Ministry of Science and ICT 2018). Under him, ITU-T did pursue some of those controversial areas (perhaps reflecting ITU membership's drive), but he also fostered dialogues about including more privacy considerations. Post-2022, a South Korean, Lee Jae-hee, was elected Deputy Secretary-General of ITU – indicating SK's engagement in steering ITU. With a foot in both worlds, South Korea can bring Western concerns to ITU privately in Asian diplomatic style, possibly softening the confrontation.

South Korea also can influence G7/G20 indirectly. It's not a G7 member, but in 2021 the UK invited South Korea to the G7 summit as a guest (along with India and Australia) and those guests signed onto a *"Open Societies Statement"* that included commitments to human rights online and free flow of information (G7 UK Cornwall 2021). That statement, while not AI-specific, underscores a shared front on digital governance values with Western allies. In G20, South Korea usually aligns with consensus endorsing principles like the 2019 AI Principles. South Korea's presence in these gatherings ensures that Asia is represented on the side of rights-respecting governance – undermining any narrative that AI ethics is "West vs rest".

It's worth noting Japan as well (though not one of the case's main trio), as Japan is another middle-power bridging actor. Japan co-led the Osaka Track (G20 data governance) and

insisted on human-centric AI in G20 2019. Japan and South Korea often coordinate on tech standards in Asia forums like APEC, offering an alternative to China's approach. So, collectively, these large middle powers, sometimes called the *"like-minded"* or *"value-sharing"* democracies, act as a network that propagates the normative framework beyond Euro-Atlantic contexts.

Impact of Middle-Power Agency

The involvement of India, Australia, and South Korea has generally *reinforced the normative convergence among democracies and at the same time, extended the geographic legitimacy* of those norms. When principles like human rights in AI are championed not just by Western powers but also by an Asian giant (India) and others, it's harder for skeptics to dismiss them as Western imposition. These countries also adapt norms to local contexts, improving their practicality. For instance, India pushed for *transparency and explainability in AI in local languages* (ensuring AI ethics isn't limited to English-speaking contexts) – such inputs can influence ISO guidance on user information or UNESCO's recommendation to include cultural and linguistic diversity as part of fairness (Effoduh 2024; Radiya-Dixit 2024).

Large middle powers also sometimes act as a *balancing voice*. If the US and EU were at odds (like possibly on how strict to regulate), they can mediate. This ultimately supports polycentric governance where multiple voices co-create norms. That said, these powers have their own interests: India and South Korea maintain significant trade with China, so they are careful not to appear overtly anti-China in multilateral settings. They thus often frame AI norms in positive terms (what they support) rather than directly criticising Chinese practices in those forums. But behind scenes, they coordinate with Western allies to ensure the outcome aligns with democratic values. Australia is less hesitant to call out China, especially on issues like

surveillance tech exports in the Pacific (they actively provide alternatives to Pacific nations so they don't adopt Chinese systems blindly). This shows varying styles: Australia more openly offensive norm diffusion (like EU/US), India/SK more bridging/diplomatic, yet all toward embedding a rights-respecting orientation globally.

Therefore, these large middle powers significantly contribute to a *polycentric convergence model*: they bring additional centres of norm diffusion that collaborate with the main Western centres but also locally contextualise them. They exemplify the bridging strategy in competitive norm diffusion – they are not normatively neutral, but they often can talk to both sides. For example, Indian PM Modi at global fora will speak of the need for international cooperation on AI with trust (resonating with Western discourse) and simultaneously stress sovereignty (which China nods to). This rhetorical mix can create space for consensus language.

As a result, fundamental rights get articulated not only in Western philosophical terms but also as universal and developmental values – strengthening their position. For example, non-discrimination in AI is framed by these states as not just a human right but as essential to social harmony in diverse societies (India, SK multiethnic to degrees), making it a universally relatable goal. Privacy is advanced as necessary for citizen trust in technology, which even developing country governments now see as key to digital adoption.

By analyzing these cases, we have seen various facets of the geopolitics of AI standards: normative export (EU), alliance alignment (US-EU), contestation (China vs democracies at ITU), and bridging (large middle powers). The next section synthesizes these findings to discuss the overarching "sovereignty–interoperability paradox" – that tension between national control and global alignment – and how the competitive norm diffusion model helps explain the current trajectory of AI standardisation governance.

CHAPTER VI: SOVEREIGNTY–INTEROPERABILITY PARADOX IN COMPETITIVE NORM DIFFUSION

As evidenced above, the evolution of AI standardisation is marked by a fundamental tension: the desire of states (and blocs) to assert *digital sovereignty* over AI governance versus the need for *interoperable standards and frameworks* that allow AI systems to function and be trusted across borders. This chapter now synthesises insights from the cases to explicate this sovereignty–interoperability paradox, using the competitive norm diffusion model and typology (offensive, defensive, bridging strategies) as an analytic lens. It discusses how each major actor's approach to AI standards reflects an attempt to reconcile, or at least manage, this paradox, and assesses whether the net outcome is tending toward a polycentric convergence (multiple actors aligning on baseline standards) or a duo-centric fragmentation (splitting into incompatible spheres).

Digital sovereignty in AI context means a state's ability to control and govern AI within its jurisdiction according to its own laws, norms, and interests – e.g., setting its own rules for data usage, algorithmic accountability, or even developing indigenous AI standards to reduce dependency on others. *Interoperability* refers to the capacity of AI systems, data, and frameworks to work seamlessly across jurisdictions – which typically requires common technical standards, shared vocabularies, mutual recognition of compliance regimes, and broadly compatible normative frameworks (so that what is considered safe and trustworthy in one region is also seen as such in another). Sovereignty emphasises divergence (each region doing things its own way), whereas interoperability necessitates convergence (agreeing on some unified approaches). In the realm of fundamental rights, this paradox is acute: different political systems have different interpretations of rights (e.g., the West prioritizes individual privacy,

China prioritizes state security or collective order), yet AI is a globally traded and distributed technology, so completely divergent standards (say, one that ignores privacy vs one that mandates strong privacy) will create friction, mistrust in cross-border AI systems, and possibly technological decoupling.

EU's Assertive Convergence

The EU's strategy can be seen as attempting to resolve the paradox by *exporting its sovereignty* – i.e., turning its internal standards into global ones so that by asserting its normative sovereignty (via the AI Act and ethics requirements) it actually creates a form of interoperability (since others adopt those standards). This is classic *Brussels Effect* logic. Rather than accept a patchwork, the EU bet that its head start in regulation would lead firms and other countries to align to its rules, thereby harmonising standards around EU values. In practice, we see this working to an extent: ISO 42001 aligns with EU-friendly principles, and the US even adjusted its approach to incorporate EU-like principles (through NIST, etc.). However, the EU's approach can also inadvertently create fragmentation if others do not follow. For instance, if China or smaller states decided EU's requirements are too onerous and thus they prefer alternative standards, that would result in dueling standards regimes. The EU mitigates this by working within international bodies to get buy-in (e.g., making sure ISO standards support compliance with the AI Act) (O'Brien, Rasdale, and Wong 2024). This is essentially an *offensive norm diffusion* strategy that seeks to achieve interoperability by persuading (or pressuring via market access) others to adopt the same norms – achieving convergence by expanding the EU's regulatory sphere outward.

One can frame it as EU pursuing *sovereignty through interoperability*: by making its rules the basis for international standards, it maintains control (sovereignty) over outcomes yet

gains the benefits of global alignment (interoperability). But this works primarily among like-minded countries or where market incentives are strong. With China, this strategy hits limits – China might partially adopt (rhetorically) but maintain different practices. That sets a boundary on how global the Brussels Effect can go; indeed, analysts predict the AI Act *"will have global impact, but a limited Brussels Effect"*, expecting that democracies and multinational companies will adapt, but China and other authoritarian regimes may not.

US's Voluntary Alignment

The US historically prioritised its tech industry's freedom (digital laissez-faire), reflecting digital sovereignty in the sense of *not binding itself to international rules* that might hinder US companies. Yet as AI risks and global pressures mounted, the US found that completely divergent approaches could hurt its own interests – for instance, if European and Asian markets demand trustworthy AI and US firms don't deliver, they lose out. Hence the US pursued interoperability via soft governance: creating voluntary frameworks (NIST RMF) that could be globally referenced and collaborating with the EU to ensure compatibility. This is a less assertive approach than the EU's but aims to meet in the middle. It is effectively *bridging from a position of power*: the US didn't want to surrender sovereignty by adopting EU law, but through bilateral work, it ensures the core concepts align so that a company can comply with both with minimal friction (NIST 2022). We might call this *cooperative norm diffusion*: rather than one side dominating, the US and EU diffused norms to each other and outward together (e.g., joint promotion of OECD standards) (Esteves and Klingebiel 2021).

However, the US still hasn't committed to hard international rules. So in a sense, it preserves sovereignty (no binding external constraints) but leverages interoperability where beneficial. This carefully managed balance shows in the TTC statements about *"ensure the*

interoperability of AI regulations" – which implicitly acknowledges separate regulations but pledges to make them work together. The US and EU, therefore, are striving for a *duo-centric cooperative convergence* – two poles working to avoid divergence between them.

China's Dual Track Strategy

China's approach to sovereignty is explicit: it emphasizes "cyber sovereignty" and has erected its own frameworks (Great Firewall, unique regulations like its 2022 algorithm regulations, 2021 Personal Information Protection Law that differs from GDPR in giving state more leeway). In standards, China's *Standardization Strategy (China Standards 2035)* envisions leading or creating standards that other countries follow, thus expanding its tech ecosystem's reach (Teleanu 2021). This is an offensive diffusion attempt to create *interoperability on its own terms*, perhaps one day a parallel global network where Chinese standards are the norm among participating countries (some talk of a "Digital Silk Road standards sphere"). The *ITU case* shows this attempted in microcosm.

At the same time, China cannot fully detach from global standards that facilitate trade (like core ISO/IEC standards for product safety, etc.). In AI, Chinese companies want to sell abroad, so they do conform to certain international standards for interoperability (e.g., Chinese firms actively contribute to ISO SC 42 foundational standards on AI vocabulary and system engineering – that indicates they accept some level of common language). But where standards implicate political control or sensitive tech (like surveillance or Internet architecture), China will push its sovereign view, interoperability be damned. This dual track – conform internationally in generic areas, diverge or lead in strategically important areas – is likely to continue. It might result in *fragmentation in some domains but not others*. For example, we

might see global alignment on AI safety testing methods (less politically charged), but split standards on things like data governance or biometric surveillance (politically charged).

China's signing of broad principles (OECD, UNESCO) could be seen as a bridging gesture, but constructivist scholars might call it *"norm shaping"* – join consensus to avoid isolation, but interpret norms in its own way when implementing. Indeed, the UNESCO AI ethics recommendation, while agreed by China, is non-binding; concurrently, China enforces AI use in ways arguably inconsistent with those principles (e.g., censorship AI harming freedom of expression). So in effect, China supports interoperability in rhetoric (to not be seen as rogue) but maintains sovereignty in practice. This posture creates a *façade of convergence hiding real divergence*. Over time, if Chinese tech continues to spread without adopting democratic norms, we could have a bipolar tech environment: devices and systems from China that embed one set of assumptions (like built-in content filtering, facial recognition with no privacy constraints) and Western devices that embed another (privacy by design, etc.). Interoperability at purely technical levels might still exist (they might interconnect data formats), but *normative interoperability* (users trusting systems from the other sphere) will suffer. We already see glimpses: Western nations distrust Chinese AI systems for security/rights reasons (e.g., US bans on Chinese drones citing data privacy), while China distrusts Western AI for ideological reasons (it blocked ChatGPT fearing uncensored info). Those are normative incompatibilities manifesting as separate markets.

Adaptive Multilateralism of Large Middle Powers

India, Australia, South Korea, etc., navigate between these poles. They cherish aspects of sovereignty – e.g., India insists on data sovereignty (mandating local storage of certain data) and has not signed some Western-led data agreements – but they also champion global

interoperability in principles (India wants globally applicable AI ethics that also account for developing world needs). Their bridging role helps smooth paradoxical tensions: by contributing to global standards in SC 42 and GPAI, they shape interoperability frameworks while safeguarding their interests (e.g., ensuring development concerns are integrated means they don't feel their sovereignty compromised by one-size-fits-all rules). They might be inclined to follow Western standards but with some localisation. For example, if an AI standard includes human rights, India will likely adopt but interpret "non-discrimination" to possibly also cover caste or other local protected classes; that's an adaptation that doesn't break interoperability but retains cultural specificity.

Large middle powers also engage in *forum shopping for cooperation*, they often convene or join forums specifically to build consensus (like GPAI, Quad) that can counter fragmentation by presenting united fronts across regions (Quad 2021a). This fosters polycentric governance through adaptive multilaterism – multiple hubs (North America, Europe, Asia-Pacific democracies, etc.) but linked by common values and agreements. Arguably, this polycentric network is a strategy to ensure interoperability among democracies while each maintains enough flexibility domestically (no single hegemon among them – more distributed leadership). It stands in contrast to a bipolar split (democratic vs authoritarian bloc), because large middle powers engage non-aligned states too, trying to pull them towards the shared standards orbit. For instance, if India persuades some African or ASEAN partners to adopt GPAI's best practices or EU-like laws, that expands the interoperable sphere of liberal norms, constraining the rival sphere.

CONCLUSION

In this thesis, I set out to examine how the United States, European Union, China, and large middle powers (India, Australia, South Korea) contest, negotiate, and diffuse AI standards in key international forums, and what impact these dynamics have on fundamental rights such as privacy, non-discrimination, freedom of expression, and due process. I sought whether the trajectory of AI governance is towards a convergent, polycentric order or a fragmented, bipolar one. The analysis across chapters yields clear answers:

1. **Major Power Strategies:** The US, EU, and China each leveraged international bodies to project their preferred AI governance models. The EU pursued an *"offensive norm entrepreneurship"* – successfully promoting its *trustworthy AI* framework globally, influencing ISO standards and even China's public rhetoric. The US, after initial hesitation, converged with the EU on core principles, using *liberal-institutionalist cooperation* (e.g., TTC) to align risk management approaches and endorse international standards supporting rights. China engaged in *forum shopping* at the ITU and elsewhere to embed its state-centric, surveillance-oriented norms, but encountered stiff resistance when those clashed with liberal values. Thus, contestation was intense: Western democracies formed a fairly unified front to ensure AI standards reflect democratic values, while China pushed back by advancing parallel standards and emphasizing sovereignty.
2. **Middle-Power Agency:** India, Australia, and South Korea emerge as pivotal *norm bridges*. They actively participated in shaping standards and normative initiatives (GPAI, Quad) to ensure broader adoption of human-rights-aligned principles beyond the West. For example, India's leadership in GPAI infused global AI discussions with

concerns of inclusivity and equity, bridging North-South perspectives. Australia amplified democratic norms in the Indo-Pacific and within technical committees, reinforcing alignment among allies. South Korea contributed technical expertise and moderated between US and Asian views, as seen in its dual roles in OECD and ITU. These large middle powers did not passively choose between US/EU and China – they largely sided with the democratic normative camp, but also shaped it by adding their regional viewpoints, thereby *diffusing norms in a culturally adaptive manner*. Their role in forum orchestration (like Quad's principles or pushing OECD norms in G20) was crucial to building a global coalition for trustworthy AI that extends well beyond North America and Europe.

3. **Impact on Fundamental Rights:** The case studies demonstrate that the dynamics in standard-setting directly affect how well fundamental rights are safeguarded in AI governance. Where EU and allied influence dominated (ISO 42001, NIST RMF), standards include systematic consideration of privacy, fairness, and accountability. Conversely, the Chinese-led standard proposals (ITU facial recognition) glaringly omitted or undermined those rights. The proactive stance of the democratic coalition prevented the adoption of standards that would normalize rights violations. Additionally, through bodies like UNESCO and OECD, fundamental rights principles have been formally affirmed as central to AI governance by a majority of nations (House 2022). This normative entrenchment means that any international AI initiative now faces the expectation to address human rights – a significant development compared to a decade ago. That said, enforcement and actual practice lag behind: the EU AI Act will operationalize rights protections in one bloc, but globally there's still a gap between principle and practice. Nevertheless, the trend is that fundamental rights have become a

key battleground and a key benchmark in AI standardisation geopolitics – with democracies largely succeeding in putting them on the agenda of every serious standards discussion.

The contest over AI standards is a microcosm of a larger contest: whether digital technologies will be governed in line with democratic values or authoritarian ones. In this contest, the past five years show that democratic coalitions, leveraging both market power and moral suasion, have made significant strides in embedding fundamental rights into the lingua franca of AI governance. However, it is an ongoing process requiring vigilance. If we lapse, authoritarian regimes could still codify their norms and export them via technology, undermining rights globally. The competitive norm diffusion model articulated here highlights that this is not a static or one-sided process – it is dynamic, with actors adapting and sometimes norms crossing divides (e.g., China adapting "trustworthy AI" language is a form of norm diffusion, albeit superficial so far) (Siegmann and Anderljung 2022). Ultimately, while full convergence on values may be elusive, a pragmatic goal is to establish *guardrails* in global AI standards that protect core human dignity and freedoms no matter who's AI it is or where it's deployed. The geopolitics of AI standardisation will undoubtedly intensify as AI becomes even more central to economies and societies; but so too can the resolve to ensure that our fundamental rights, painstakingly codified over decades, are not lost in the algorithms that govern our future. The work must continue – in standards meetings, diplomatic negotiations, technical research, and public discourse – to steer AI towards enhancing, not eroding, the rights and freedoms that define our shared humanity.

BIBLIOGRAPHY

Angst, Mario, Jack Mewhirter, Danielle McLaughlin, and Manuel Fischer. 2022. 'Who Joins a Forum—And Who Does Not?—Evaluating Drivers of Forum Participation in Polycentric Governance Systems'. *Public Administration Review* 82 (4): 692–707. https://doi.org/10.1111/puar.13427.

Austin, Marc Theodore. 2000. 'Competing for Global Standards: The Political Economy of International Standard Setting in High Technology Industries - ProQuest'. 2000. https://www.proquest.com/openview/38af3fda7bad45d251adf2cdd89ebf16/1?pq-origsite=gscholar&cbl=18750&diss=y.

Australian Human Rights Commission. 2021. 'Final Report: Human Rights and Technology | Australian Human Rights Commission'. 1 March 2021. https://humanrights.gov.au/our-work/technology-and-human-rights/publications/final-report-human-rights-and-technology.

Bazoobandi, Sara, Sangeeta Mahapatra, Tetiana Schipper, and Iris Wieczorek. 2025. 'From Global Governance to Nationalism: The Future of AI'. 2025. https://www.giga-hamburg.de/en/publications/giga-focus/from-global-governance-to-nationalism-the-future-of-ai.

'Beijing Artificial Intelligence Principles'. 2019. International Research Center for AI Ethics and Governance. 2019. https://ai-ethics-and-governance.institute/beijing-artificial-intelligence-principles/.

Blind, Knut. 2009. 'Standardisation as a Catalyst for Innovation'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. https://papers.ssrn.com/abstract=1527333.

Bradford, Anu. 2020. 'The Brussels Effect: How the European Union Rules the World'. *Faculty Books*, March. https://doi.org/10.1093/oso/9780190088583.001.0001.

Burt, Chris. 2019. 'Standards for Biometric Surveillance Being Drafted for ITU by Chinese Businesses | Biometric Update'. 2 December 2019. https://www.biometricupdate.com/201912/standards-for-biometric-surveillance-being-drafted-for-itu-by-chinese-businesses.

Busch, Lawrence. 2011. *Standards: Recipes for Reality*. The MIT Press. https://doi.org/10.7551/mitpress/8962.001.0001.

Cameran, Ashraf. 2020. 'Artificial Intelligence and the Rights to Assembly and Association'. *Journal of Cyber Policy* 5 (2): 163–79. https://doi.org/10.1080/23738871.2020.1778760.

Cameran, Ashraf. 2022. 'Exploring the Impacts of Artificial Intelligence on Freedom of Religion or Belief Online'. *The International Journal of Human Rights* 26 (5): 757–91. https://doi.org/10.1080/13642987.2021.1968376.

Cameran Hooshang, Ashraf. 2015. 'The Spatiality of Power in Internet Control and Cyberwar'. UCLA. https://escholarship.org/uc/item/0w99g31p.

Chen, Zhe, Chuang Wang, Guanwen Li, Zhe Lou, Sheng Jiang, and Alex Galis. 2020. 'NEW IP Framework and Protocol for Future Applications'. In *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, 1–5. https://doi.org/10.1109/NOMS47738.2020.9110352.

China Academy of Information and Communications Technology. 2021. 'White Paper on Trustworthy Artificial Intelligence'. *Center for Security and Emerging Technology* (blog). July 2021. https://cset.georgetown.edu/publication/white-paper-on-trustworthy-artificial-intelligence/.

Cihon, Peter. 2019. 'Standards for AI Governance: International Standards to Enable Global Coordination in AI Research & Development | GovAI'. 17 April 2019. https://www.governance.ai/research-paper/standards-for-ai-governance-international-standards-to-enable-global-coordination-in-ai-research-development.

Collins, Aengus, Marie-Valentine Florin, Anca Rusu, Gianluigi Viscusi, Gianluca Misuraca, and Andrea Renda. 2023. 'Dissemination Level: Public', May.

Comerma, Dr Laia. 2024. 'Postdoctoral Researcher under the Hans van Baalen Scholarship', no. 26.

Crawford, Kate. 2021. *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. https://doi.org/10.2307/j.ctv1ghv45t.

Dawson, Dave, Emma Schleiger, Joanna Horton, McLaughlin, and Cathy Robinson. 2019. 'Artificial Intelligence: Australia's Ethics Framework - a Discussion Paper'. 5 April 2019. https://apo.org.au/node/229596.

DeNardis, Laura. 2009. *Protocol Politics: The Globalization of Internet Governance*. The MIT Press. https://doi.org/10.7551/mitpress/9780262042574.001.0001.

Derivry. 2023. '[ARTICLE] Digital Sovereignty in India: Policy Agenda, Discourse, Power and Capability'. 4 May 2023. https://webserver07.reims.sciences-po.fr/public/chaire-numerique/en/2023/05/04/contribution-digital-sovereignty-in-india-policy-agenda-discourse-power-and-capability/.

Directorate-General for Communication. 2024. 'Commission Launches AI Innovation Package'. Text. European Commission - European Commission. 24 January 2024. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383.

Effoduh, Jake Okechukwu. 2024. 'A Global South Perspective on Explainable AI'. Carnegie Endowment for International Peace. 30 April 2024. https://carnegieendowment.org/research/2024/04/a-global-south-perspective-on-explainable-ai?lang=en.

Egan, Michelle. 2002. 'Setting Standards: Strategic Advantages in International Trade'. *Business Strategy Review* 13 (1): 51–64. https://doi.org/10.1111/1467-8616.00202.

Ekdal, Dino. 2021. 'Normative Power Europe & AI'. Lund University.

Engel, Annegret, and Xavier Groussot. 2025. *The EU's Digital Package: Striking a Balance for Fundamental Rights in the DSA and DMA Regulations*. *Swedish Studies in European Law*. Vol. 19. Hart Publishing Ltd. http://lup.lub.lu.se/record/4a8eb403-19d4-47ae-b015-50cba4f6b5e1.

Engler, Alex. 2023. 'The EU and U.S. Diverge on AI Regulation: A Transatlantic Comparison and Steps to Alignment'. *Brookings* (blog). 25 April 2023. https://www.brookings.edu/articles/the-eu-and-us-diverge-on-ai-regulation-a-transatlantic-comparison-and-steps-to-alignment/.

Esteves, Paulo, and Stephan Klingebiel. 2021. 'Diffusion, Fusion, and Confusion: Development Cooperation in a Multiplex World Order'. In *The Palgrave Handbook of Development Cooperation for Achieving the 2030 Agenda: Contested Collaboration*, edited by Sachin Chaturvedi, Heiner Janus, Stephan Klingebiel, Xiaoyun Li, André de Mello e Souza, Elizabeth Sidiropoulos, and Dorothea Wehrmann, 185–215. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-57938-8_9.

EU Commission. 2020. 'White Paper on Trustworthy Artificial Intelligence'. 19 February 2020. https://cset.georgetown.edu/publication/white-paper-on-trustworthy-artificial-intelligence/.

'European Union: Initiatives on AI and Emerging Technologies US-EU TTC'. 2022. 2022. https://digitalpolicyalert.org.

Feakin, Tobias. 2025. 'AI Geopolitics Beyond the US-China Rivalry'. *Aspen Digital* (blog). 7 March 2025. https://www.aspendigital.org/blog/ai-geopolitics-beyond-the-us-china-rivalry/.

Feldstein, Steven. 2024. 'Evaluating Europe's Push to Enact AI Regulations: How Will This Influence Global Norms?' *Democratization* 31 (5): 1049–66. https://doi.org/10.1080/13510347.2023.2196068.

Ford, The Hon Christopher. 2022. 'CHINA'S STRATEGIC VISION | PART THREE ENVISIONING A SINOCENTRIC WORLD'. *OCCASIONAL PAPERS* 1 (1).

G7 UK Cornwall. 2021. '2021 Open Societies Statement'. 13 June 2021. https://g7.utoronto.ca/summit/2021cornwall/210613-open-societies.html.

G7, United Kingdom. 2021. 'G7 Digital and Technology Track - Annex 1'.

G20 Delhi Declaration. 2023. 'G20 Digital Economy Ministers Meeting Outcome Document and Chair Summary'. 2023. https://g7g20-documents.org/database/document/2023-g20-india-sherpa-track-digital-economy-ministers-ministers-language-g20-digital-economy-ministers-meeting-outcome-document-and-chair-summary.

G20 Digital Economy Ministers Meeting. 2020. 'Ministerial Declaration: G20 Digital Economy Ministers Meeting, July 22, 2020'. 22 July 2020. https://www.g20.utoronto.ca/2020/2020-g20-digital-0722.html.

Gabbott, Miranda. 2023. 'The Bumpy Road Toward Global AI Governance', September. https://www.noemamag.com/the-bumpy-road-toward-global-ai-governance.

Gaur, Akriti. 2025. 'Cross-Border Data Flows and India's Digital Sovereignty'. *Verfassungsblog*, March. https://doi.org/10.59704/82a368a8a6158abf.

Ginsburg, Tom. 2020. 'How Authoritarians Use International Law'. *Journal of Democracy* 31 (4): 44–58. https://muse.jhu.edu/pub/1/article/766183.

Haigh, Edoardo Campanella, John. 2025. 'China Wants to Run Your Internet'. *Foreign Policy* (blog). 4 June 2025. https://foreignpolicy.com/2023/08/25/china-wants-to-run-your-internet/.

Hao, Karen. 2020. 'The US Just Released 10 Principles That It Hopes Will Make AI Safer'. MIT Technology Review. 7 January 2020. https://www.technologyreview.com/2020/01/07/130997/ai-regulatory-principles-us-white-house-american-ai-initiatve/.

Heeks, Richard, Ospina ,Angelica V., Foster ,Christopher, Gao ,Ping, Han ,Xia, Jepson ,Nicholas, Schindler ,Seth, and Qingna and Zhou. 2024. 'China's Digital Expansion in the Global South: Systematic Literature Review and Future Research Agenda'. *The Information Society* 40 (2): 69–95. https://doi.org/10.1080/01972243.2024.2315875.

High-Level Expert Group on AI. 2019. 'Ethics Guidelines for Trustworthy AI | Shaping Europe's Digital Future'. 8 April 2019. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

Horowitz, Michael C. 2018. 'Artificial Intelligence, International Competition, and the Balance of Power (May 2018)', May. http://hdl.handle.net/2152/65638.

House, Chatham. 2022. 'Reclaiming Human Rights in a Changing World Order | 8. Technical Standards and Human Rights: The Case of New IP'. 10 October 2022. https://www.chathamhouse.org/2022/10/reclaiming-human-rights-changing-world-order/8-technical-standards-and-human-rights-case.

IISS. 2019. 'Australia, Huawei and 5G'. IISS. 2019. https://www.iiss.org/ja-JP/publications/strategic-comments/2019/australia-huawei-and-5g/.

Ingersleben-Seip, Nora von. 2023. 'Competition and Cooperation in Artificial Intelligence Standard Setting: Explaining Emergent Patterns'. *Review of Policy Research* 40 (5): 781–810. https://doi.org/10.1111/ropr.12538.

'ISO/IEC JTC 1/SC 42 - Artificial Intelligence'. 2017. Participating Members. 2017. https://www.iso.org/committee/6794475.html?view=participation.

Jili. 2022. ‘The Rise of Chinese Surveillance Technology in Africa (Part 1 of 6)’. *EPIC - Electronic Privacy Information Center* (blog). 31 May 2022. https://epic.org/the-rise-of-chinese-surveillance-technology-in-africa/.

Kania, Elsa B., and John and Costello. 2021. ‘Seizing the Commanding Heights: The PLA Strategic Support Force in Chinese Military Power’. *Journal of Strategic Studies* 44 (2): 218–64. https://doi.org/10.1080/01402390.2020.1747444.

Kareem, Karwan. 2025. *The Cyber Eye: Inside the Network Wars and Secrets of the Five Eyes Alliance*. Lulu Press, Inc. https://open.icm.edu.pl/handle/123456789/25802.

Kulothungan, Vikram, Priya Ranjani Mohan, and Deepti Gupta. 2025. ‘AI Regulation and Capitalist Growth: Balancing Innovation, Ethics, and Global Governance’. In *2025 IEEE 11th Conference on Big Data Security on Cloud (BigDataSecurity)*, 39–45. https://doi.org/10.1109/BigDataSecurity66063.2025.00020.

Latonero, Mark. 2018. ‘Governing Artificial Intelligence’: Data & Society. https://datasociety.net/wp-content/uploads/2018/10/DataSociety_Governing_Artificial_Intelligence_Upholding_Human_Rights.pdf.

Lebryk, Theo. 2021. ‘China Focus | The Fight over the Fate of the Internet: The Economic, Political, and Security Costs of China’s Digital Standards Strategy’. *China Focus* (blog). 21 April 2021. https://chinafocus.ucsd.edu/2021/04/21/the-fight-over-the-fate-of-the-internet-the-economic-political-and-security-costs-of-chinas-digital-standards-strategy/.

Lott, Philip. 2022. ‘9DASHLINE — How China Became the Standard Maker’. 9DASHLINE. 26 October 2022. https://www.9dashline.com/article/how-china-became-the-standard-maker.

Matthews, William. 2025. ‘The World Should Take the Prospect of Chinese Tech Dominance Seriously, and Start Preparing Now | Chatham House – International Affairs Think Tank’. 29 January 2025. https://www.chathamhouse.org/2025/01/world-should-take-prospect-chinese-tech-dominance-seriously-and-start-preparing-now.

Mattli, Walter. 2001. ‘The Politics and Economics of International Institutional Standards Setting: An Introduction’. *Journal of European Public Policy* 8 (3): 328–44. https://doi.org/10.1080/13501760110056004.

Mattli, Walter, and Tim Büthe. 2003. ‘Setting International Standards: Technological Rationality or Primacy of Power?’ *World Politics* 56 (1): 1–42. https://doi.org/10.1353/wp.2004.0006.

McGeachy, Hilary. 2019. ‘US-CHINA TECHNOLOGY COMPETITION: IMPACTING A RULES-BASED ORDER’. *May 2019*, May.

McInerney, Kerry. 2024. 'Yellow Techno-Peril: The "Clash of Civilizations" and Anti-Chinese Racial Rhetoric in the US–China AI Arms Race - Kerry McInerney, 2024'. 4 June 2024. https://journals.sagepub.com/doi/full/10.1177/20539517241227873.

Melson, Richard. 2021. 'China Shaping UN Facial Recognition & Surveillance Standards'. 14 January 2021. https://metaintelligence.org/world-watching-chinese-tech-groups-shaping-un-facial-recognition-and-surveillance-standards/.

Ministry of Science and ICT. 2018. 'Press Releases - 과학기술정보통신부 >'. 2018. https://www.msit.go.kr/eng/bbs/view.do?sCode=eng&nttSeqNo=314&pageIndex=&searchTxt=&searchOpt=&bbsSeqNo=42&mId=4&mPid=2.

Mueller, Alex, and Christopher S. Yoo. 2023. 'Crouching Tiger, Hidden Agenda?: The Emergence of China in the Global Internet Standard-Setting Arena'. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4528546.

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

National Security Commission on Artificial. 2021. 'Final Report: National Security Commission on Artificial Intelligence'. Report. UNT Digital Library. National Security Commission on Artificial Intelligence (U.S.). United States. 1 March 2021. https://digital.library.unt.edu/ark:/67531/metadc1851188/.

NIST. 2022. 'TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management', December.

NITI Aayog. 2018. 'National Strategy for Artificial Intelligence', June.

O'Brien, Claire, Mark Rasdale, and Daisy Wong. 2024. 'The role of harmonised standards as tools for AI act compliance | DLA Piper'. 11 January 2024. https://www.dlapiper.com/es-pr/insights/publications/2024/01/the-role-of-harmonised-standards-as-tools-for-ai-act-compliance.

OECD. 2019. 'AI Principles'. OECD. 2019. https://www.oecd.org/en/topics/ai-principles.html.

OECD. 2021. 'State of Implementation of the OECD AI Principles: Insights from National AI Policies'. OECD Digital Economy Papers 311. Vol. 311. OECD Digital Economy Papers. https://doi.org/10.1787/1cd40c44-en.

Olson, Stephen. 2020a. 'China's Influence to Strengthen'. Hinrich Foundation. 28 April 2020. https://www.hinrichfoundation.com/research/article/tech/china-influence-strengthens/.

Olson, Stephen. 2020b. 'China's Influence to Strengthen'. Hinrich Foundation. 28 April 2020. https://www.hinrichfoundation.com/research/article/tech/china-influence-strengthens/.

OSTP. 2022. 'Blueprint for an AI Bill of Rights | OSTP'. *The White House* (blog). 2022. https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/.

Ostrom, Elinor. 2010. 'Beyond Markets and States: Polycentric Governance of Complex Economic Systems - American Economic Association'. June 2010. https://www.aeaweb.org/articles?id=10.1257/aer.100.3.641.

Panday, Jyoti. 2025. 'Sovereign Data Strategies: Boosting or Hindering AI Development in India?' Orfonline.Org. 10 January 2025. https://www.orfonline.org/expert-speak/sovereign-data-strategies-boosting-or-hindering-ai-development-in-india.

Park, Min Chul, and Seong-Hyeon Bang. 2020. 'Legislative Trends in AI: National AI Ethical Standards and Amendment to Enforcement Decree of Framework Act on Intelligent Informatization - Kim & Chang'. 2020. https://www.kimchang.com/en/insights/detail.kc?sch_section=4&idx=22510.

PIB. 2022. 'India Takes over as Council Chair of Global Partnership on AI (GPAI)'. 21 November 2022. https://www.pib.gov.in/PressReleasePage.aspx?PRID=1877739.

Prabhakaran, Vinodkumar, Margaret Mitchell, Timnit Gebru, and Iason Gabriel. 2022. 'A Human Rights-Based Approach to Responsible AI'. arXiv. https://doi.org/10.48550/arXiv.2210.02667.

PTI. 2022. 'India to Take over Chair of Global Partnership on AI for 2022-23'. 21 November 2022. https://cio.economictimes.indiatimes.com/news/government-policy/india-to-take-over-chair-of-global-partnership-on-ai-for-2022-23/95661329.

Quad. 2021a. 'Joint Statement from Quad Leaders'. *The White House* (blog). 25 September 2021. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2021/09/24/joint-statement-from-quad-leaders/.

Quad. 2021b. 'Quad Principles on Technology Design, Development, Governance, and Use'. 2021. https://openresearch-repository.anu.edu.au/items/a9c8f550-46cc-470a-9774-05d7a6ee3786.

Radiya-Dixit, Evani. 2024. 'Local AI Research Groups Are Preserving Non-English Languages in the Digital Age | TechPolicy.Press'. Tech Policy Press. 21 October 2024. https://techpolicy.press/local-ai-research-groups-are-preserving-nonenglish-languages-in-the-digital-age.

Regulation (EU) 2024/1689. 2024. http://data.europa.eu/eli/reg/2024/1689/oj/eng.

Ricciardi Celsi, Lorenzo, and Albert Y. Zomaya. 2025. 'Perspectives on Managing AI Ethics in the Digital Age'. *Information* 16 (4): 318. https://doi.org/10.3390/info16040318.

Schaefer, Brett D., and Danielle Pletka. 2022. 'Countering China's Growing Influence at the International Telecommunication Union'. 7 March 2022. https://www.geopolitic.ro/wp-content/uploads/2022/03/BG3689.pdf.

Schmitt, Lewin. 2022. 'Mapping Global AI Governance: A Nascent Regime in a Fragmented Landscape'. *AI and Ethics* 2 (2): 303–14. https://doi.org/10.1007/s43681-021-00083-y.

Shaelou, Stéphanie Laulhé, and Yulia Razmetaeva. 2023. 'Challenges to Fundamental Human Rights in the Age of Artificial Intelligence Systems: Shaping the Digital Legal Order While Upholding Rule of Law Principles and European Values'. *ERA Forum* 24 (4): 567–87. https://doi.org/10.1007/s12027-023-00777-2.

Sherman, Justin. 2022. 'China's War for Control of Global Internet Governance'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.4174453.

Siegmann, Charlotte, and Markus Anderljung. 2022. 'The Brussels Effect and Artificial Intelligence | GovAI'. August 2022. https://www.governance.ai/research-paper/brussels-effect-ai.

Sinha, Amber. 2024. 'The Landscape of Facial Recognition Technologies in India'. 13 March 2024. https://www.techpolicy.press/the-landscape-of-facial-recognition-technologies-in-india/.

Sukumar, Arun, and Arindrajit and Basu. 2024. 'Back to the Territorial State: China and Russia's Use of UN Cybercrime Negotiations to Challenge the Liberal Cyber Order'. *Journal of Cyber Policy* 9 (2): 256–87. https://doi.org/10.1080/23738871.2024.2436591.

Tabassi, Elham. 2023. 'Artificial Intelligence Risk Management Framework (AI RMF 1.0)'. NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology (U.S.). https://doi.org/10.6028/NIST.AI.100-1.

Tassey, Gregory. 2000. 'Standardization in Technology-Based Markets'. *Research Policy* 29 (4): 587–602. https://doi.org/10.1016/S0048-7333(99)00091-8.

Techatassanasoontorn, Angsana A., and Shuguang Suo. 2011. 'Influences on Standards Adoption in de Facto Standardization'. *Information Technology and Management* 12 (4): 357–85. https://doi.org/10.1007/s10799-011-0089-2.

Teleanu, Sorina. 2021. 'The Geopolitics of Digital Standards: China's Role in Standard-Setting Organisations', December.

The White House, The White. 2021. 'Joint Statement from Quad Leaders'. *The White House* (blog). 25 September 2021. https://bidenwhitehouse.archives.gov/briefing-room/statements-releases/2021/09/24/joint-statement-from-quad-leaders/.

'TTC Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management'. 2022, December.

UNESCO. 2022. 'Recommendation on the Ethics of Artificial Intelligence'. 2022. https://unesdoc.unesco.org/ark:/48223/pf0000381137.

Vijayakumar, Anupama. 2024. 'AI Ethics for the Global South: Perspectives, Practicalities, and India's Role'. ResearchGate. December 2024.

https://www.researchgate.net/publication/388198144_AI_Ethics_for_the_Global_South_Perspectives_Practicalities_and_India's_role.

Villarino, José-Miguel Bello y. 2023. 'Global Standard-Setting for Artificial Intelligence: Para-Regulating International Law for AI?', October. https://doi.org/10.1163/26660229-04101018.

Warner, Jason, and Toyosi Ajibade. 2024. 'China's Smart Cities in Africa: Should the United States Be Concerned?', November. https://www.csis.org/analysis/chinas-smart-cities-africa-should-united-states-be-concerned.

Wouters, Jan. 2023. 'Corporations and the Making of Public Standards in International Law: The Case of China in the International Telecommunication Union'. In *The Evolution of Transnational Rule-Makers through Crises*, edited by M. Konrad Borowicz, Panagiotis Delimatsis, and Stephanie Bijlmakers, 66–82. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009329408.005.

Yin, Jessie. 2024. 'The Race for Cyberspace: China's IP Standards and the Threat to Net Neutrality'. *Chinaobservers* (blog). 10 October 2024. https://chinaobservers.eu/the-race-for-cyberspace-chinas-ip-standards-and-the-threat-to-net-neutrality/.