

Charting the Geopolitics and European Governance of Artificial Intelligence

Raluca Csernatoni

Charting the Geopolitics and European Governance of Artificial Intelligence

Raluca Csernatoni

© 2024 Carnegie Endowment for International Peace. All rights reserved.

Carnegie does not take institutional positions on public policy issues; the views represented herein are those of the author(s) and do not necessarily reflect the views of Carnegie, its staff, or its trustees.

No part of this publication may be reproduced or transmitted in any form or by any means without permission in writing from Carnegie Europe or the Carnegie Endowment for International Peace. Please direct inquiries to:

Carnegie Endowment for International Peace Publications Department 1779 Massachusetts Avenue NW Washington, D.C. 20036 P: + 1 202 483 7600 F: + 1 202 483 1840 CarnegieEndowment.org

Carnegie Europe Rue du Congrès, 15 1000 Brussels, Belgium P: +32 2 735 56 50 Carnegie Europe.eu

This publication can be downloaded at no cost at CarnegieEurope.eu.

Contents

Summary	1
Introduction	3
Scripting the Future: Probing Narratives of Al Power and Disruption	5
The EU at the Helm? Navigating AI Geopolitics and Governance	9
To Innovate or to Regulate? That Is the Misleading Question	16
Conclusion	22
About the Author	25
Notes	27
Carnegie Europe	33

Summary

Artificial intelligence (AI) stands out as a transformative force that is shaping global power dynamics, security considerations, and economic paradigms. With general-purpose AI systems like OpenAI's ChatGPT, the AI revolution has transitioned from a future prospect to a present-day reality that attracts significant attention from great powers and tech giants alike. However, alongside its promises, AI also causes skepticism when it comes to the future-proofing governance frameworks, broader ethical considerations, and the technology's dual-use implications in the civilian and military domains.

Against this backdrop, the European Union (EU) governs AI in multifaceted ways while navigating geopolitical, economic, and regulatory concerns. A nuanced understanding is needed of the EU's AI technopolitics and the ways this is reflected in European efforts to govern this field, foster AI innovation, and ensure trustworthiness. In this regard, the EU's AI Act aims to regulate the use of AI systems based on risk levels, reflecting a commitment to the human-centric and responsible development of AI. Moreover, the EU's pursuit of homegrown AI innovation underscores AI's critical importance in bolstering European technological sovereignty and reducing strategic dependencies.

On the one hand, the EU views establishing a global standard through the AI Act as a pivotal objective, prompting discussions of a regulatory race. Such debates point not only to the significance of AI regulation but also to the potential for the EU to wield considerable influence globally. Yet, amid the global race for AI supremacy, Europe faces challenges in setting a gold standard for AI regulation and maintaining a technological edge. While specific provisions of the AI Act may exert substantial influence on global markets, Europe's efforts alone will not establish a comprehensive international standard for AI.

On the other hand, while the EU has allocated substantial investment through various programs, competition from other major economies, particularly the United States and China, remains formidable. In response, the EU will need to match its rhetoric of technological sovereignty on AI with significant funding. As things stand, there is little evidence that the EU will be able to pursue technological sovereignty and global leadership in the AI domain given Europe's lack of major high-tech companies and investment.

Narratives of AI power and AI disruption further complicate the EU's technopolitical landscape. The key questions in both sets of narratives are whom exactly they serve and what technopolitical realities they shape. These narratives range from framing AI as a national security concern to seeing the technology as a disruptive force with the potential to trigger paradigm shifts in the civilian and military sectors. Such narratives play a central role in defining policy and governance agendas for future technological progress and global reordering. Often amplified by powerful actors, they structure perceptions of AI's transformative potential, from utopian visions of economic growth to dystopian fears of existential risks with catastrophic effects.

Addressing these prevailing narratives is crucial to establishing realistic governance expectations and fostering a balanced public discourse in Europe. Debunking exaggerated claims and focusing on tangible risks and challenges can guide EU policy actions, innovation efforts, and the operationalization of the AI Act in alignment with broader societal concerns and values.

Introduction

The fast-paced evolution of emerging and disruptive technologies (EDTs), such as artificial intelligence (AI), heralds novel forms of power, digital disruption, and economic influence on the global stage. With the increased use of AI systems, especially general-purpose AI like ChatGPT, there is a widespread perception of a rapidly unfolding AI revolution. AI is the new holy grail for companies involved in the technology's development—primarily the field's leaders, OpenAI and Google subsidiary DeepMind—not least given the vast profits and the major recognition that would come with being the first company to develop human-level machine intelligence. The rapid acceleration of AI is occurring concurrently with a resurgence of great-power politics, a deepening crisis in the multilateral global order, and the growing role of corporate tech giants in international relations. So-called weaponized interdependence is also proliferating across critical economic nodes, technological supply chains, and global information flows.¹

At the same time, there is rising popular skepticism about the ability of governments and international organizations like the European Union (EU) to empower citizens to reap perceived opportunities while protecting people from the dangers of this technology. Against this backdrop, prevailing narratives of power and disruption are shaping the trajectories of technological progress and global reordering. Debunking such narratives, whether they are true or false, is imperative in the context of EU governance and innovation, because they often carry considerable influence as well as mold perceptions, policy decisions, and societal expectations.

For instance, according to Carnegie's Matt O'Shaughnessy, the current "hype over AI superintelligence could lead policy astray," distract attention from more pressing risks and

challenges, divert resources, and, consequently, shape policy actions in line with the interests of powerful actors.² Exposing such approaches helps establish realistic governance expectations and contributes to a more balanced discourse on the role of AI in collective futures. The same can be said of the EU's rhetoric of technological sovereignty. Despite the EU's push for technological sovereignty and aspirations of global AI leadership, there is a notable gap between Europe's ambition and its achievement. Notwithstanding protectionist tendencies and claims of technological superiority, the EU faces challenges because of its relatively limited digital industry and its modest investments compared with those of great powers like the United States and China.

That is why a more nuanced understanding of the international and European technopolitics governing AI systems is needed. In this context, the term "technopolitics" refers to the complex interplay between AI and the geopolitical structures, power dynamics, narratives, norms, and economic influences that shape and are shaped by AI technologies. This analysis zooms in on the EU's position in the geopolitics of AI, the current proliferation of overlapping global governance and regulatory frameworks, the web of narratives surrounding AI, the EU's funding and innovation initiatives, and broader ethical considerations. The goal is to examine the interplay between AI, regulation, and geopolitics in the EU while emphasizing the sociopolitical and normative implications of AI research, development, and use in Europe.

In the arena of AI governance, the EU faces a formidable challenge amid intensifying state and corporate rivalry alongside the emergence of a complex web of regulatory regimes worldwide. The EU's dedication to fostering responsible AI practices is laudable. Yet, the bloc's ability to forge a unified foreign policy stance on AI, cultivate strategic partnerships with major allies, implement the EU's AI Act effectively across the member states, and navigate the myriad of international governance initiatives will be pivotal in shaping AI's global trajectory. As Europe seeks to assert itself as a leading force in AI governance, these strategic imperatives will serve as key indicators of the union's technological power, global influence, and efficacy in navigating the increasingly competitive landscape of AI regulation and innovation.

The technopolitics of AI becomes evident through narratives of the technology's power and disruption. These narratives are not merely rhetorical tools wielded by powerful actors but integral components of world politics and the broader technopolitical and social landscape. The crucial concerns with both types of narrative are which actors they serve and what technopolitical realities they shape. The dominance of such narratives, which cast AI as a harbinger of both power and disruption, has significantly influenced decisionmaking processes, policy agendas, investment priorities, and public perceptions. While not entirely incorrect, these narratives can be misleading in their portrayal of AI's potential and overly restrictive in defining the range of possible AI futures. Rooted in aspirations of global leadership, military prowess, and economic dominance, these narratives often overlook nuanced considerations of ethics, norms, values, rules, governance, and societal impact. As AI technologies continue to evolve, it becomes imperative to critically evaluate these narratives.

This analysis views AI systems not as neutral but as a social construct influenced by the discourses, visions, values, societal norms, economic and security interests, and power structures of the technology's creators, funders, deployers, and users. The notion of AI technopolitics acknowledges that the creation, benefits, and control of AI are not evenly distributed. Certain regions and actors, such as governments, institutions, large corporations, and expert communities, wield significant influence in shaping the narratives and, consequently, the trajectories of AI's development and governance.

In essence, the concept of AI technopolitics recognizes that the technology is not merely a technical phenomenon but a sociotechnical one: decisions about design, deployment, and governance foreground power dynamics and have far-reaching implications for geopolitical rivalries, governance structures, social relations, and the distribution of benefits and risks in society. Therefore, to better understand the EU's position, it is crucial to untangle the rhetoric surrounding AI's geopolitical and transformative potential and the way this potential impacts global and European power structures and economic paradigms.

Scripting the Future: Probing Narratives of AI Power and Disruption

Narratives of Al's power and disruption share a common thread: the seeming inevitability of the technology's transformative impact.³ First, these narratives maintain that it is unavoidable to imagine a future in which AI power is not translated into technological arms races, contests over global leadership, and struggles for economic influence among states and corporate players. Second, it is argued, an AI-induced disruption is an all-pervasive and inevitable transformation, which may go as far as to lend agency to AI models that transcend human control. Yet, the key questions with both types of narrative are whom they serve and what technopolitical realities and public perceptions they shape.

Narratives of Al Power

Narratives of power influence critical decisions about which AI applications will receive funding, regulatory guidelines, or societal acceptance. They also set expectations of what role AI will play in enhancing geopolitical power in a new arms race to harness AI as a critical strategic enabler for national and European security. There are important sociopolitical implications to the claim that AI is a crucial instrument of statecraft, power projection, and economic prowess that will irrevocably affect the rise and fall of polities.⁴

The EU is no exception in this regard. With its European AI Strategy, the EU aims to strengthen its technological sovereignty and become a world-class hub for AI, both by trying to build strategic leadership in this sector and by ensuring that AI is human-centric and trustworthy.⁵ Such ambitious goals have been translated into a specific European approach to fostering excellence and trust in this domain, from boosting the innovation, development, and uptake of AI across the bloc to putting forward a harmonized legal framework for trustworthy AI, as set out by the AI Act.6

However, amid the discourse surrounding the EU's technological sovereignty and a surge of initiatives aimed at positioning the EU as a leader in AI, a closer examination reveals a significant disparity between rhetoric and reality. Despite the EU's concerted efforts to assert its AI sovereignty through various strategies, including protectionist measures and aspirations of technological superiority, the path to true leadership is fraught with challenges. With a limited digital tech industry and relatively low investment compared with industry giants like the United States and China, the EU's ambitions of technological sovereignty and AI leadership face considerable hurdles.

The AI Act, on which the European Parliament and the EU Council reached an agreement in December 2023 after lengthy negotiations, also faces obstacles. The act is the world's first-ever attempt to legislate for a comprehensive, horizontal, and risk-based regulation of AI systems in use. With its narrative of trustworthy AI, the EU is defining its vision of an AI-disrupted future while aiming to consolidate its role as a normative power in the field of AI governance. Yet, the EU will need to grapple with several challenges in its quest to establish the bloc as a benchmark for AI regulation and retain its technological primacy. While certain aspects of the act may have considerable sway over global markets, Europe's regulatory efforts alone are insufficient to set forth a comprehensive global standard for AI.

The EU's efforts to establish technological sovereignty, especially in the realm of AI systems, are intrinsically linked to the bloc's pursuit of strategic autonomy in defense. For instance, as the EU seeks to enhance its security and defense capabilities, the European Defence Fund supports actions that can help develop EDTs based on concepts that originate from "non-traditional defence's state-of-art" or emerging technologies like AI systems. Such technologies are equally disruptive in the civilian and the military domain; in the latter, they contribute to the development of innovative defense systems.

Indeed, the ongoing war in Ukraine has become an AI war lab for tech giants. Companies including Palantir, Microsoft, Amazon, Google, Clearview AI, and others have collaborated with Ukrainian armed forces to provide advanced technologies and support for Ukraine against Russia's aggression. This collaboration has turned Ukraine into a laboratory for testing military technologies, including AI, drones, and facial-recognition software. This close partnership between defense forces and tech giants aims not only to assist in the immediate conflict but also to develop Ukraine's tech sector for long-term economic growth.

Yet, the increasing experimentation with military AI technologies highlights the complexities and ethical concerns surrounding the use of these technologies in wartime as well as the potential global implications of such collaboration. One immediate takeaway is that

the theater of conflict serves as a crucible for unfolding AI power in the realm of warfare. Against this backdrop, tech companies emerge as independent actors that wield outsize power, while military decisions are likely to be handed off to algorithms.

In the global race for AI supremacy, all major economies, including the EU, are fervently striving to secure their technological power and leading position.¹⁰ The United States asserts its dominance by leading the world with the highest total private investment in the field. In 2022, the United States secured \$47.4 billion in AI investment, approximately 3.5 times more than the next-highest contributor, China, which reported \$13.4 billion. The United States is also home to the largest number of newly funded AI companies, surpassing the combined total of the EU and the United Kingdom almost twofold and outpacing China with 3.4 times as many enterprises.¹¹

Europe, despite its scientific excellence, finds itself trailing behind, especially on the financing front. To compete on the global stage, emerging startups require substantial financial backing, and there is a pressing need to fortify and interconnect high-tech ecosystems across Europe to effectively transform innovative ideas from the lab to the market and, importantly, into global commercial success. The European Commission has taken several measures to advance AI technologies: the Horizon 2020 program allocated €1.5 billion (\$1.6 billion) to AI in 2018-2020; the Digital Europe Program, as part of the EU's 2021-2027 multiyear budget, dedicated an additional €2.5 billion (\$2.7 billion) to investing in and opening up the use of AI by businesses and public administrations.¹²

Nevertheless, the question is whether this financing will be enough to preserve Europe's technological power, sovereignty, and economic security. The EU needs to contend with formidable corporate players and states that are leveraging huge amounts of funding and jostling for a first-mover position on both the innovation and the regulation of AI. On the regulatory front, building trustworthy AI and addressing the risks generated by specific uses of AI may prove equally challenging. With the AI Act, the EU aims to address such risks through a set of complementary, proportionate, and flexible rules. The legal framework set out by the act is based on four levels of risk and introduces dedicated rules for general-purpose AI models. However, AI systems are continually evolving and may amplify risks or challenges in unforeseen ways, such as exacerbating social inequalities, exposing security vulnerabilities, endangering defense applications, and restricting access to technological benefits.

For the AI Act to become an influential global rule book, the EU will need to successfully operationalize and future-proof the legislation once it enters into force—expected by 2026—by making sure that the act effectively mitigates potential future harms and disparities. Yet, it remains to be seen whether these rules will also provide Europe with a leading role as a powerful global actor in setting an international regulatory gold standard.

Narratives of Al Disruption

At the same time, AI systems are often framed as disruptive technologies. The commission refers to a disruptive technology, including AI, as one that induces

"a disruption or a paradigm shift, i.e. a radical rather than an incremental change. Development of such a technology is 'high risk, high potential impact', and the concept applies equally to the civil, defence and space sectors. Disruptive technologies for defence can be based on concepts or ideas originating from non-traditional defence actors and find their origins in spin-ins from the civil domain."13

This rather general definition focuses on dual-use, enhanced or new AI technologies that bring about a radical change, including a paradigm shift in the concept and conduct of civilian and military affairs, like replacing existing technologies or rendering them obsolete.

This definition leaves ample room to ask what AI-triggered disruption means, what the effects are of powerful actors heralding this disruption, and what the broader sociopolitical contexts are within which disruptive technologies emerge. There is also the important question of how to separate fact from fiction in the context of often-hyped AI developments. Narratives of AI disruption are intimately linked to narratives of geopolitical and economic power on the global stage. In the case of AI, "the confusion over what disruption means, who exercises it, and upon whom is not a coincidence: rather, disruption's polysemy [multiple meanings] is structurally produced as a way to disguise ongoing capitalist crisis as a technical problem that market innovations can solve," in the words of political scientist Charmaine Chua.¹⁴ Narratives of technological disruption have become powerful instruments to cement depictions of either attainable, desirable futures of economic growth and societal progress or dystopian renderings of catastrophic risks. These narratives also prescribe the kinds of speculative futures that should be attained by AI-led technological innovation and progress in both the civilian and the military domain.

The notion of disruption implies a preexisting, linear, temporal dimension that is meant to be disrupted, and the mere claim of disruption is a performative act that can—but does not always—trigger, enable, or enact preimagined sociotechnical utopian or dystopian futures.¹⁵ For instance, it is now uncontroversial to think that an AI-induced extinction of humanity is a genuine possibility. In terms of dystopian narratives of disruption, the existential risk of artificial general intelligence (AGI) is a powerful and radical idea—that substantial progress in this field could result in human extinction or an irreversible global catastrophe on a par with nuclear annihilation or a pandemic. There are also growing fears and superstitions stemming from the belief that sci-fi-esque AI superintelligence will soon emerge and take over humanity. 16 The first step to demystifying such popular and contested narratives is to collect more concrete evidence about the underlying claim that AI will represent an existential risk. Yet, it is equally important to recognize these narratives' power of persuasion and explore their origins and whom they serve.

European Commission President Ursula von der Leyen adopted a similar narrative in her 2023 State of the Union address, when she quoted from a statement by the Center for AI Safety: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war." At the same time, von der Leyen underlined the narrative of power, noting that "AI is a general technology that is accessible, powerful and adaptable for a vast range of uses—both civilian and military."

The EU at the Helm? Navigating AI **Geopolitics and Governance**

The past year has witnessed unprecedented advancements in the capabilities of AI systems, coupled with numerous international, European, national, and civil society initiatives to govern and regulate this field. The year 2023 was marked by significant progress in establishing AI governance frameworks as well as intense competition and strategic positioning among leading nations and international organizations, particularly the United States and China, and, to a certain extent, the EU.

This geopolitical rivalry extends beyond technological competition and also encompasses norms promotion, standard setting, the security of supply chains, economic security, and even ideologically driven liberal and illiberal narratives about the future trajectories of AI. Advanced militaries are actively pursuing sophisticated and dual-use AI technologies while experimenting with, prototyping, and deploying increasingly autonomous functions across the battlefield and in the cyber-physical domain. 18 The dynamics between AI innovation, regulation, and geopolitics are thus a feature that defines and structures world politics.

For the EU to be at the helm of international AI governance would signify a concerted effort to shape the global norms, standards, and regulations that govern the development and deployment of AI technologies. Such a position would reflect the EU's ambition to promote a human-centric and trustworthy approach to AI. This goal aligns with the EU's broader objective of fostering technological sovereignty and AI innovation while ensuring the protection of fundamental rights and values. Achieving a leadership position would entail actively participating in multilateral forums, collaborating with partners and other stakeholders, and advancing proposals for harmonized regulatory frameworks that address the complexities and challenges posed by AI technologies.

However, in realizing this ambition, the EU faces several obstacles, not least the diversity of interests among its member states and the fragmented nature of its decisionmaking processes. While the EU has made strides in advancing AI governance through initiatives such as the AI Act, the union's influence on the global stage remains limited compared with that of dominant players like the United States and China. Moreover, the lack of a unified approach among EU member states can hinder the bloc's ability to speak with a coherent voice on the global stage.

The Geopolitics of AI and Technological Power

At the international level, the geopolitics of AI unfolds as a high-stakes rivalry among great powers, notably the United States and China, engaged in a complex race to establish supremacy in developing AI capabilities and setting international norms. This rivalry has far-reaching implications for Europe's economic aspirations, military power, and global influence. The United States, home to leading tech giants and cutting-edge research institutions, leverages its technological innovation to maintain a global competitive edge. Conversely, China, propelled by state-backed initiatives and a massive pool of data, seeks to assert its dominance in AI applications, posing a formidable challenge to Western leadership. China is also leading the way in AI regulation by publishing pioneering strategies to govern AI systems.¹⁹

Increasingly, both states view AI as a crucial element of national security, having integrated it into their security and defense strategies and raised concerns about an AI arms race and the weaponization of AI.²⁰ For instance, in August 2023, the U.S. Department of Defense launched a new task force to investigate the possibility of using generative AI (GenAI), such as large language models (LLMs), for military missions.²¹ Yet, while the United States and China are global AI rivals, they are also the most important AI collaborators. According to a 2023 report by Stanford University, the number of AI research collaborations between the two competitors quadrupled between 2010 and 2021, although the rate of collaboration has since slowed significantly and will likely continue to do so because of national security concerns.²²

Whoever is winning the AI arms race, however, is not the key issue. Rather, the mere perception of an arms race may push governments and tech giants to eschew trustworthy and responsible AI and cut corners in safety research and regulation. This could be disastrously risky. Besides, fears that illiberal and autocratic China might catch up with Western liberal-democratic contenders may jeopardize efforts to endorse AI governance and regulatory initiatives that could slow down dangerous developments. Equally, the perception of an AI arms race may impede the creation of a global AI governance framework.

Despite not being able to leverage the same scale of tech giants or state-backed initiatives as the United States and China, the EU has used its research base, as well as its regulatory and market power, to govern AI systems responsibly according to the rule of law and democratic oversight.²³ While AI has become an important "tool of power politics" and of national security for individual states, the EU has approached the technology "primarily from an economic, social, and regulatory" angle, in the words of political scientist Ulrike Franke.²⁴ At the same time, the EU has increasingly adopted the language of power and a geopolitical stance toward critical dual-use technologies like AI. According to European Commissioner for Internal Market Thierry Breton, mastery of technology is central to the "new geopolitical order."25 As early as 2019, at the start of her mandate at the helm of what she called a geopolitical commission, von der Leyen noted, "First, we must have mastery and ownership of key technologies in Europe. These include quantum computing, artificial intelligence, blockchain, and critical chip technologies."26

Gaining this mastery is, of course, easier said than done. Several challenges are noteworthy, particularly the EU's unique and hybrid polity, which requires better coordination between and across EU institutions and member states. Crafting a common EU foreign policy and security agenda on AI is another priority area for EU and national decisionmakers. Likewise, addressing the dominance of U.S.-driven AI innovation and research in Europe is a key concern for the EU's strategic autonomy, technological sovereignty, and economic security to prevent dependencies.²⁷

There is also a risk of growing nationalist tendencies and protectionist state actions that can further accelerate the geopolitical race for AI and deepen digital divides.²⁸ The EU's own efforts to build up its digital and technological sovereignty mean that it, too, risks falling into the protectionist trap. Besides pursuing a deeper strategic partnership with the United States, the EU will need to cultivate other geopolitical contenders and potential partners, such as India, Israel, Japan, South Korea, Taiwan, and the United Arab Emirates, that have meaningful roles to play in shaping the future of AI.

Two further important questions are what AI means for the concentration of corporate hegemony and what role tech giants play in world politics. Before 2014, when Google acquired DeepMind, most notable machine-learning models were released by academia; however, a noteworthy shift has taken place since, with industry assuming a dominant role. In 2022, the landscape displayed a stark asymmetry, with thirty-two major machine-learning models originating from the industry, while academia contributed merely three.²⁹ This transformation underscores the evolving dynamics of the field, in which corporate entities have become the primary drivers of influential innovation. This trend includes the monopolization of critical infrastructure connectivity nodes, such as satellites and undersea cables; the commodification of big data and data extraction; the leveraging of colossal budgets for AI innovation; the headhunting of leading AI talent; and the control and amassing of impressive compute capacity.

The United States stands as the EU's closest ally and strategic partner, and transatlantic collaboration on trustworthy AI and other crucial governance matters is essential in confronting challenges posed by fast-evolving AI models.³⁰ However, framing U.S.-European cooperation on AI solely in the context of countering China is not the most prudent approach. While there is a shifting sentiment in Europe regarding China, it is significant that Europeans do not necessarily share the U.S. sense of urgency in pushing back against China. Consequently, the U.S. aspiration to leverage transatlantic AI cooperation as a strategy to curb Chinese influence might have limited resonance in Europe. This has been evidenced by French President Emmanuel Macron's call for Europe to reduce its reliance on the United States and steer clear of entanglements in the increasing tensions between Washington and Beijing.³¹ Implicit in this vision is an aspiration for European strategic autonomy and a desire for the EU, possibly under France's leadership, to emerge as a distinct and influential third superpower on the global stage.

Initiatives like the EU-U.S. Trade and Technology Council (TTC) and the Global Partnership on Artificial Intelligence (GPAI)—an international, multistakeholder initiative to guide the responsible development and use of AI consistent with human rights, fundamental freedoms, and shared democratic values—highlight the shared commitment to responsible AI.³² In May 2023, after the fourth meeting of the TTC, European Commission Executive Vice President Margrethe Vestager revealed a collaborative EU-U.S. effort to formulate an AI code of conduct to be presented to Group of Seven (G7) leaders, with companies encouraged to participate voluntarily. The objective of establishing nonbinding international standards that encompass risk audits, transparency, and other requirements for AI companies was reached in October 2023, when G7 leaders agreed on the International Guiding Principles for AI and the Code of Conduct for AI developers.³³

The Race to AI Governance and Regulation

The race to innovate on AI has also triggered a race to regulate.³⁴ The increased visibility of the technology's risks and challenges has led to calls, especially in the EU, for policymakers, regulators, and civil society organizations to look beyond the opportunities and benefits by putting forward governance and regulatory frameworks that ensure AI is human-centric, responsible, safe, and trustworthy. With the surge of GenAI systems like ChatGPT, regulating the use of such systems in society, the economy, and the workplace is crucial to mitigate unforeseen or undesirable consequences.

Hence, the question of how to ensure effective and adequate governance of AI has become front and center in global and domestic debates. Besides minimizing risks and challenges, such governance and regulatory frameworks could enable better and faster uptake of AI in the public and the private sector, boost legal certainty, and, therefore, contribute to advancing economies' positioning in the global race. Over the past few years, several international, European, and national governance and regulatory frameworks for AI have emerged, alongside the creation of high-level expert and advisory groups as well as various multilateral, minilateral, and bilateral forums.

Given the absence of a single, comprehensive global governance structure for AI, this recent proliferation of initiatives can be understood as a regime complex—a network or system of interconnected international agreements, institutions, and norms that govern a particular issue or area. Indeed, the global and cross-border nature of AI technologies and AI-related challenges, coupled with the interconnectedness of economies, calls for multilateral and multistakeholder collaboration to develop shared principles and standards.

This backdrop should mandate a collective approach to establishing an effective and adaptive global governance framework. However, the current landscape of initiatives presents a fragmented and overlapping regime complex that reflects the varying needs, interests, and values of different regions, governments, and stakeholders. This emerging competition to set global norms will be difficult for the EU to navigate, not least because the bloc currently relies on

what law professor and Carnegie nonresident scholar Anu Bradford has called the "Brussels effect"—the EU's unilateral power to regulate global markets.³⁵ In the case of AI, this approach "aims to enact effective domestic regulations on AI developments, and then rely on the direct or extraterritorial effects of such regulations to affect the conditions or standards for AI governance in other jurisdictions," in the words of researchers Matthijs Maas and José Jaime Villalobos.36

Yet, will the EU's AI Act have a Brussels effect and become the global gold standard for AI regulation? EU legislators often invoke this effect as a driving force behind the act's fast ratification. The EU's rush to approve the legislation has been fueled not only by the act's primary objectives of safeguarding European consumers and, to a lesser degree, fostering AI innovation in the bloc: it has also been driven by a fear that Europe risks forfeiting its edge in digital governance as rival nations bolster their regulatory capabilities, narrowing the gap with the EU. This erosion of the union's comparative advantage threatens to diminish the EU's pivotal role as a primary driver of global regulatory standards.

A variety of multilateral organizations have published principles for AI governance, such as the Organisation for Economic Co-operation and Development's (OECD's) Principles on Artificial Intelligence; the GPAI's principles for responsible stewardship of trustworthy AI; the United Nations Educational, Scientific, and Cultural Organization's Recommendations on the Ethics of Artificial Intelligence; and the EU's Ethics Guidelines for Trustworthy AI.³⁷ The Council of Europe, an international human rights body, set up the Committee on Artificial Intelligence at the beginning of 2022 to develop a Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law.³⁸

Tech giants have also put forward various AI principles, such as Microsoft Azure's AI Principles, which offer a guide for the development and application of AI in the company, and Google's Ethical AI Principles, which serve as a framework for evaluating new AI products and features. Other examples include Amazon's commitment to the responsible use of AI technologies and OpenAI's approach to AI safety. Meanwhile, the World Economic Forum's Global AI Governance Alliance is an initiative that unites industry leaders, governments, academic institutions, and civil society "to champion responsible global design and release of transparent and inclusive AI systems."39

To delineate the corporate ethical AI agenda, three broad regulatory strategies are possible: first, an absence of legal regulation, with ethical principles and responsible practices relegated to voluntary and nonbinding commitments; second, a middle ground involving soft regulatory frameworks that do not substantially conflict with innovation and profitability; and third, hard regulation that restricts or prohibits the deployment of the technology.

Predictably, the tech sector leans toward the first two options and resists the third. This preference is exemplified by the Tech Accord to Combat Deceptive Use of AI in 2024 Elections, which was signed by twenty companies, including Adobe, Amazon, Google, IBM, Meta, Microsoft, OpenAI, TikTok, and X, and announced during the 2024 Munich Security

Conference. 40 While it is promising to see such companies acknowledge the wide-ranging harms posed by AI, the principles proposed under the accord are generic and reactive and do not proactively address the potential weaponization of content that is deceptively fake or alters the appearance, voice, or actions of political figures during elections. The accord's commitments are declaratory and lack nuance in terms of defining harmful AI-generated content, disinformation, and weaponization.

At the national level, according to Stanford University's 2023 report, countries passed thirty-seven AI-related bills in 2022, and as of 2023, more than sixty countries had published national AI strategies.⁴¹ Most of these regulations not only call for better analysis and understanding of AI and its potential risks and benefits but also demand that AI developers be made accountable for the actions of their inventions.

Indeed, in the United States several guidance documents and voluntary frameworks have emerged in the past few years, such as the AI Risk Management Framework of the U.S. National Institute of Standards and Technology (NIST), published in January 2023, and the White House's Blueprint for an AI Bill of Rights, a set of high-level principles issued in October 2023. 42 In February 2024, the administration of U.S. President Joe Biden named Elizabeth Kelly, a top driving force behind the president's October 2023 executive order on the development and use of AI, as the director of the new U.S. AI Safety Institute (USAISI) at NIST.43

The administration also announced the creation of the U.S. AI Safety Institute Consortium, to be hosted under the USAISI. This consortium will bring together AI developers, researchers, government officials, industry experts, and civil society representatives in support of advancing the development and implementation of safe and trustworthy AI. Aligning with Biden's landmark executive order, the consortium will play a pivotal role in taking priority actions, such as devising guidelines for red teaming, conducting capability assessments, managing risk, ensuring safety and security, and implementing measures to authenticate synthetic content.

Meanwhile, China has demonstrated its active engagement in formulating principles and regulations to govern AI, from the 2017 New-Generation AI Development Plan and expert-driven initiatives such as the 2019 Governance Principles for New-Generation AI to, most notably, the 2023 Measures for the Management of Generative AI Services.⁴⁴ As for the EU, despite the innovative nature of the AI Act, it remains to be seen how the union will navigate the interplay between voluntary principles, codes of conduct, soft laws, norms, and binding regulations.

The development of GenAI has led to an array of new initiatives, including the OECD's 2023 G7 Hiroshima Process on Generative AI.45 Shortly after the G7 agreed on its AI principles and code of conduct, the United Kingdom in November 2023 hosted the AI Safety Summit, which gathered senior government officials, executives of major AI companies, and civil society leaders to lay the foundations for a global AI safety regime.⁴⁶ The result was the

Bletchley Declaration, a joint commitment by twenty-eight governments and leading AI companies that called for advanced AI models to undergo safety checks before deployment.⁴⁷

Both the G7 commitments and the Bletchley Declaration recognized the risks of advanced foundation models and set out voluntary best practices for the industry. In October 2023, the UN announced the creation of an AI Advisory Body to issue guidance on the risks, opportunities, and international governance of AI.⁴⁸ It remains to be seen how this thirty-nine-member body will support the international community's efforts to govern AI and build a global scientific consensus on risks and challenges. In a surprisingly fast development, the advisory body launched an interim report on governing AI as early as December 2023.⁴⁹

While the EU had a head start and a first-mover advantage in setting the global agenda with its AI Act as a blueprint for other governments, the regime complex highlights an increasingly crowded AI global governance landscape that will be difficult to navigate. It has been reported that the European Commission is preparing to push back against a U.S.-led attempt to exempt the private sector from the Council of Europe's work to develop a framework convention for AI.⁵⁰ The EU's goal is to create as much alignment as possible with the AI Act, but there are pending issues related to the scope of the convention, such as whether to exempt companies by default, with the option for signatories to involve the private sector voluntarily, and what a national security exemption should look like. In this respect, the commission is in favor of clearly excluding from the convention AI systems developed solely for national security, military, and defense purposes, in line with the AI Act and the positions of EU member states such as France.

The global AI regime complex also indicates the emergence of an international epistemic community for sharing narratives of safe, trustworthy, and responsible AI. According to researchers from the University of California and Princeton University, this community "is sustained through its mutually reinforcing community-building and knowledge production practices" in various online and offline forums.⁵¹

The EU and its member states should take an active part in, or at least keep a close eye on, such initiatives to ensure that these forums are not watered down, avoid divisions among member states, and strengthen the EU's common position on responsible AI and AI safety. Any effective EU AI policy toward the United States and China will also depend on a common EU foreign policy on AI, which currently seems to be lacking.

Finally, any global approach to AI governance that omits China is poised to yield only marginal results, whereas the inclusion of China would invariably reshape the agenda. The latter scenario would introduce into debates a myriad of contentious issues to do with democracy and human rights, thus expanding the scope of discussions and altering the prevailing narrative to reflect China's distinct perspective on AI governance and policies.

To Innovate or to Regulate? That Is the **Misleading Question**

The EU finds itself at the crossroads of AI innovation and AI regulation. A tension arises from the challenge of fostering innovation to maintain global competitiveness on AI while establishing robust regulatory frameworks that address concerns related to fundamental rights, democratic principles, and societal impact. Striking the right balance is crucial for the EU to harness AI's transformative potential while upholding the union's commitment to human-centric and responsible development of AI.

The European Commission has emerged as a prolific policy entrepreneur that has put forward numerous governance measures to mitigate this dilemma. These include policy initiatives and projects to create expert groups, financing platforms for industry consortia, public-private partnerships on AI-enabling technologies, and a proposed AI regulation as part of the EU's broader digital strategy. Yet, questions remain about whether such initiatives are too little too late to consolidate the EU's position in both the innovation race and the race to regulate, as well as whether a push for fast innovation can be easily married with hard regulation and strong ethical standards for the research, deployment, and use of AI.

The commission first solidified its agenda-setting role on AI in 2018 by creating the AI High-Level Expert Group (AI HLEG) to spin in and harness technoscientific, legal, and corporate expert knowledge on AI. The group gathered fifty-two experts from academia, civil society, and industry and acted as a steering body for the European AI Alliance, set up by the commission as a multistakeholder forum for dialogue on the future of AI across Europe.⁵² This expertise-building initiative underscored the commission's efforts to structure this emerging technology as a legitimate area for EU policy action. By supporting the AI HLEG with a communication on AI in April 2018 and a set of ethics guidelines for trustworthy AI in April 2019, the commission further positioned itself as the key driver of a human-centric approach to the technology.⁵³ In June 2019, the AI HLEG presented a list of thirty-three policy and investment recommendations on how to boost European industry.⁵⁴ This document was a detailed plan and vision that established the first steps to define the conditions under which AI should be developed and implemented in the EU's internal market.

In February 2020, the commission published a series of documents, including a white paper on AI, that underscored the EU's objective of becoming a front-runner in both the innovation and regulation of AI systems. According to the white paper, the EU can "become a global leader in innovation in the data economy and its applications."55 The main building blocks to achieve this goal are an "ecosystem of excellence" that mobilizes private and public sector resources along the entire value chain and an "ecosystem of trust" that ensures legal certainty for public and private sector organizations as well as rules to protect fundamental and consumer rights.

While innovation and regulation may appear to be opposing forces, the EU's approach recognizes that they are in fact complementary and mutually reinforcing. Effective regulation can provide a clear legal framework that fosters trust and confidence in AI technologies, thereby enabling their responsible adoption and widespread use. At the same time, innovation-driven policies can spur creativity and entrepreneurship, leading to the development of AI solutions that meet societal needs while complying with regulatory requirements. By striking a balance between innovation and regulation, the EU aims to create an enabling environment that promotes responsible AI innovation while safeguarding fundamental rights and values.

Safeguarding Economic Security and Innovation

When it comes to creating an AI ecosystem of excellence, an important initiative is AI Watch, which the commission launched in December 2018 as a knowledge service to monitor the development, uptake, and impact of AI across Europe.⁵⁶ AI Watch tracks the EU's industrial, technological, and research capacity in AI as well as AI-related policy initiatives in the member states. According to AI Watch, the EU invested between €13 billion and €16 billion (\$14–17 billion) in AI in 2020.⁵⁷ That year, because of the COVID-19 outbreak, the EU's AI investments grew by only around 24 percent, compared with approximately 47 percent in 2019. The EU's target is to invest €20 billion (\$22 billion) per year by 2030.⁵⁸ At the national level, France, Germany, Ireland, Italy, and Spain each invested over €1 billion (\$1.1 billion) in AI in 2020. France and Germany, the two countries that invested the most in AI, also increased their investments by more than the EU average.

In terms of Europe's AI investments relative to those of other major economies, the numbers are not encouraging. According to AI Watch, the United States spends nearly twice as much as the EU—and 2.7 times more on a per capita basis—on AI research and development and AI-related complementary assets.⁵⁹ This discrepancy raises questions about narratives of AI power and how the EU can maintain its technological edge in this field. To mitigate such gaps, the EU in 2021 outlined how the commission and the member states can harness the advantages of integrating AI into the EU's economic, societal, and environmental strategies.⁶⁰ More worryingly, in a 2022 report, the commission warned that control over technology is an increasingly crucial geopolitical battleground and that the EU is losing the investment race in various EDTs, such as quantum computing, fifth-generation (5G) technology, AI, and biotechnology.61

The commission's 2022 report further painted a bleak picture of AI investment from sources other than companies, such as venture capital and private equity: in 2015-2022, the United States captured 40 percent of all such funding globally, Europe 12 percent, and Asia (including China) 32 percent.⁶² The report did not mention EU research and development funding programs as a solution but instead proposed deepening EU banking and capital market integration to allow more private investment. The report reflected a shift at the EU level toward the prioritization of scientific links and a proactive research and innovation agenda

with like-minded democracies and partners. The report also listed key imperatives for the commission, the member states, and private stakeholders, including expediting AI investments to ensure a resilient economic and social recovery from the pandemic, implementing AI strategies and programs to maximize the EU's first-mover advantage, and harmonizing AI policies to eliminate fragmentation.

In 2020, the European Investment Bank launched a €150 million (\$162 million) facility to invest in AI alongside fund managers and private investors, while the European Investment Fund launched a pilot for an AI and blockchain investment scheme worth €100 million (\$108 million).⁶³ Beyond these funds, the commission made additional resources available under the EU's 2021–2027 budget and the post-pandemic NextGenerationEU program, specifically the Recovery and Resilience Facility, which will pay particular attention to strategic technologies like AI. In June 2023, the commission proposed a European economic security strategy, which aims to promote European competitiveness, protect the EU from economic security risks, and partner with countries that share the union's economic security concerns.64

The following month, Belgium, Finland, the Netherlands, Portugal, and Slovakia signed a joint nonpaper on the EU's open strategic autonomy. The nonpaper sought to ensure the union's capacity to cope alone if necessary but did not rule out cooperation whenever possible. The document also called on the EU to come up with a coordinated technology strategy, advance excellence in research and innovation, and speed up the transfer of technologies from the lab to scalable companies.⁶⁵ The nonpaper's framing of AI in economic security terms indicates a protectionist turn in efforts to foster an EU ecosystem of excellence: according to the document, the EU should curb certain exports to or investments in third countries while assessing the economic risks posed by the potential leakage of critical technologies amid rising geopolitical tensions with China.

In early October 2023, the commission adopted a recommendation on critical technology areas for the EU's economic security.66 Out of ten areas, the recommendation identified four that the commission considered highly likely to present the most sensitive and immediate risks related to technology security and technology leakages: advanced semiconductor technologies, AI technologies, quantum technologies, and biotechnologies. These areas were selected on the basis of the following criteria: the enabling and transformative nature of the technologies; the risk of civil-military fusion, a clear hint at China's civil-military fusion strategy; the technologies' relevance for both the civilian and the military sector and their potential to advance both domains; the risk that uses of certain technologies could undermine peace and security; and the risk that the technologies could be used in violation of human rights.

In January 2024, the commission rolled out a comprehensive AI innovation package to foster a more thriving European AI ecosystem, in which startups and innovators can work closely with industrial users, attract investment in the EU, and have access to the key

building blocks of AI, namely data, computing, algorithms, and talent.⁶⁷ This is a promising step in the right direction. The initiatives in the package also aim at boosting European startups and small- and medium-sized enterprises (SMEs) engaged in the development of trustworthy AI that would align with EU values and regulations. In this regard, von der Leyen in September 2023 unveiled a pioneering initiative to grant European AI startups access to Europe's supercomputers to enable them to train their trustworthy AI models.⁶⁸ Initiating this vision, the commission in November 2023 introduced the Large AI Grand Challenge, a competition that offers financial support and supercomputing resources to successful AI startups.

To fortify the leadership of European startups and cultivate the growth of competitive AI ecosystems across the union, the commission is set to establish what it calls "AI factories." 69 These are dynamic ecosystems that will encompass AI-dedicated supercomputers, interconnected data centers, and, critically, the skilled workforce necessary to leverage these resources effectively, from supercomputing and AI experts to data specialists, researchers, startups, and end users. The concept of AI factories will thus incorporate computing power, data, supercomputing services, and extensive initiatives to attract top-tier talent on a large scale.

Overall, in the wake of the political consensus achieved in December 2023 on the AI Act, these measures seek to advance the creation, implementation, and adoption of trustworthy AI in the EU. Yet, it remains to be seen how these initiatives will translate such commitments into tangible action, especially in terms of fostering a thriving and globally competitive cross-border AI innovation ecosystem in the EU.

Flexing the Governance Muscle With the EU's AI Act

The advent of consumer-facing AI—exemplified by the dramatic rise of ChatGPT, a generative LLM launched by OpenAI in November 2022—has brought the workings of machine-learning models into sharp focus for decisionmakers, regulators, the industry, and public opinion. This amplified attention has sparked crucial policy and regulatory concerns about safety, bias, personal data, intellectual property rights, industry dynamics, and, notably, the inherently opaque, black-box nature of the technology and the limited ability to explain and interpret its outcomes.

But GenAI is only the beginning. Even with GPT-5 set to be released by OpenAI, potentially in 2024, a team of AI scientists at Microsoft was already arguing in April 2023 that GPT-4, possibly the most sophisticated LLM yet, was "strikingly close to human-level performance" and was showing the "sparks" of AGI—a form of AI that is as smart as, or smarter than, humans in every area of intelligence, rather than simply in one task.⁷⁰

Given the potential power of new AI tools, it is sensible to place guardrails around them to minimize harm. Prominent industry leaders such as OpenAI, Google DeepMind,

Anthropic, and other leading AI labs have also sounded the alarm, suggesting that future AI systems carry the potential for catastrophic outcomes akin to pandemics and nuclear weapons.⁷¹ As the specter and hype of such existential risks loom, an urgent call for regulatory frameworks is gaining prominence, aimed at mitigating the potential dangers posed by unchecked advancements in AI. In this regard, the EU has a long history of regulating technologies that pose serious risks to public safety and human rights.

In April 2021, the commission proposed the AI Act as the EU's first regulatory framework for AI. The legislation recommends that AI systems that can be used in different applications should be analyzed and classified according to the level of risk they pose to users; these different risk levels would then be translated into more, or less, regulation.⁷² This proactive approach is crucial to avoiding undesirable outcomes and instead fostering a responsible and trustworthy AI landscape. The AI Act aims to instill trust in the offerings of AI systems for European citizens while recognizing that most systems present minimal to no risks and can actively contribute to addressing societal challenges.

Yet, there remains a critical need to address and mitigate potential high risks associated with specific AI systems. The AI Act seeks to ensure that fundamental rights, democracy, the rule of law, and environmental sustainability are protected from high-risk AI. Acknowledging the potential threats to citizens' rights and democracy posed by certain AI applications, EU legislators agreed to ban various systems, including biometric categorization that uses sensitive characteristics, untargeted scraping of facial images for recognition databases, workplace and educational emotion recognition, social scoring based on behavior or characteristics, AI systems that manipulate human behavior, and AI that exploits vulnerabilities.⁷³

For the purposes of law enforcement, safeguards and exceptions exist for biometric identification systems in public spaces, subject to judicial authorization and for defined lists of crimes. So-called post-remote biometric identification targets individuals convicted or suspected of serious crimes, while real-time biometric identification is limited to conducting targeted searches for victims, preventing terrorist threats, or localizing suspects of specific crimes.

Under the AI Act, high-risk AI systems face mandatory impact assessments for fundamental rights in sectors like insurance, healthcare, and banking. The act establishes complaint rights for citizens impacted by high-risk AI systems, especially in cases that could influence voter behavior and election outcomes. General-purpose AI systems and models must adhere to transparency requirements, including technical documentation, compliance with EU copyright law, and dissemination of training content summaries. Stricter obligations for high-impact general-purpose AI models involve evaluations, systemic risk assessments, and adversarial testing. The providers of such models must report incidents to the commission, ensure cybersecurity, and report on energy efficiency levels until harmonized EU standards are established.

In sum, the AI Act determines obligations for AI systems based on their potential risks and levels of impact. Companies that do not comply with the rules will be fined, with penalties ranging from €35 million (\$38 million) or 7 percent of a company's global annual turnover, whichever is higher, for violations of banned AI applications; €15 million (\$16 million) or 3 percent of annual turnover for violations of other obligations; and €7.5 million (\$8.1 million) or 1.5 percent for supplying incorrect information. More proportional caps are anticipated for administrative fines for SMEs and startups that breach the act's provisions.⁷⁴

Applicable in all twenty-seven EU member states, the act thus establishes a binding and novel hard regulatory framework. In terms of governance, the member states' market surveillance authorities will supervise the act's implementation at the national level. Pending approval by the European Parliament and the European Council, a new European AI Office will also be established in the commission.⁷⁵ Tasked with the enforcement and supervision of AI rules, including those governing general-purpose AI models, the office is envisioned to establish a robust connection with the scientific community and enable coordinated efforts at the European level. It will be the first body worldwide to enforce binding rules on AI and is therefore anticipated to serve as an international benchmark.

For general-purpose AI models, an essential role will be played by a panel of independent scientific experts, who will issue alerts on systemic risks and thus contribute to the classification and testing of these models. The agreement on the AI Act incorporates a tiered approach for general-purpose AI models that aims to differentiate those with potential systemic risks for society from others. The responsibility for developing methodologies and benchmarks for assessing the capabilities of such models will lie with the AI Office. Overall, the establishment of the office even before the AI Act is formally adopted demonstrates a sense of urgency. This preemptive move is driven by the need to deliver codes of conduct for disruptive general-purpose AI models within nine months of the law coming into effect. 76 Yet, resource constraints loom large. Anticipating a distribution challenge, one can speculate about which countries will contribute the most national experts to the office.

As AI is a fast-evolving technology, the question remains whether the AI Act is sufficiently future-proof and will allow its rules to adapt to technological change. AI applications should remain trustworthy even after they have been placed on the market; such a feat would require ongoing quality and risk management of the entire AI value chain. The fast-paced evolution of GenAI in 2023 took place even as the proposed act was being deliberated. The commission's initial proposal did not explicitly address this category of AI, while the council's and the parliament's proposed amendments differed substantially in their treatment of GenAI. The manifold future evolutions, applications, and uncertainties of AI systems, and their potential impacts on human rights, pose a considerable challenge to navigating the act's enforcement. As the act comes into effect, EU policymakers should carefully consider the complex relationship between the act and other EU legislation while identifying and addressing potential loopholes that may not be readily apparent and may be subtly exploited.

What is more, beyond the involvement of scientific and other expert communities or the tech sector, the inclusion of civil society in EU AI governance faces persistent and substantial barriers, such as time, financial resources, and expertise. There is also a need to translate the AI Act into accessible language that can enable engagement from those who lack AI expertise, including the broader European public. In eleventh-hour adjustments to the act, law enforcement agencies gained the authority to employ facial-recognition technology on recorded video footage without requiring judicial approval.⁷⁷ During the act's negotiations, civil society voices had warned that the use of facial-recognition technology could increase across the EU despite efforts to regulate it under the act. 78 The parliament had also called for a full ban on facial recognition but softened its redlines in response to the demands of countries such as France.⁷⁹ Paris has consistently argued that specific clauses in the act could hamper homegrown AI innovation and pose obstacles for European AI startups, including France's Mistral AI, LightOn, and Hugging Face, which aim to rival U.S. counterparts like OpenAI and Google.80

Despite the political accord reached in December 2023 on the AI Act among the EU institutions, France appears to remain steadfast in its stance and wishes to impose strict conditions to make sure that the act does not hamper the development of competitive AI models. Specifically, France argues, it is important to strike the right balance between the protection of trade secrets and transparency, avoid letting the high-risk obligations overburden companies, and reassess the thresholds and criteria used to label AI models as posing systemic risks. To clarify these finer details, on February 2, 2024, the member states' ambassadors to the EU gathered to discuss the act. France, the most reluctant country, agreed to support the text while maintaining its strict conditions after Germany and Italy lifted their reservations.⁸¹

On February 13, 2024, two of the European Parliament's committees came to a provisional agreement, by seventy-one votes to eight, on the legislation. 82 The text now awaits formal adoption in a plenary session of the parliament, expected in April 2024, and final endorsement by the EU Council. The act will be fully in force twenty-four months after it is passed, with exceptions for certain provisions.⁸³ Yet, the passing of the act marks just the beginning of its implementation journey, with potential lobbying efforts likely to persist during this phase, adding complexity to the process.

Conclusion

The year 2023 ushered in a wave of AI hype surrounding the technology's geopolitical potential and disruptive power, propelling policymakers into fervent discussions not only about the geopolitics of AI but also about safety concerns and the regulation of fast-evolving AI technologies. The GenAI summer of 2023 was characterized by unprecedented scientific advancements and economic successes in AI research, development, and deployment. Yet, amid the lofty aspirations, hyped narratives, and pledges on the geopolitics of AI, various

challenges, gaps, and inherent limitations of the technology have emerged. These call for an in-depth understanding of the influence of AI narratives on public perceptions of the technology in Europe and across the globe, especially in different cultures.

Narratives of AI's power and disruption exert a profound influence on the global stage and on the EU's strategic ambitions as a global technological power. Intertwined with geopolitical and economic power plays, these narratives shape perceptions of the future and can cause radical shifts in civilian and military affairs. Claims of AI's disruptive impacts often serve powerful actors by fostering depictions of utopian progress or dystopian risks. The discourses associated with AI can be identified as tools to control political debates and prescribe desirable or undesirable futures. In doing so, these narratives can reinforce existing power structures or create new ones, influencing the trajectory of AI's development and societal impact.

Ultimately, narratives of AI's power and disruption raise the question of who gets to imagine an AI-disrupted future—in other words, who sets the agenda for the technology's development and deployment. How stakeholders navigate such a future has become one of the main technopolitical stakes of AI systems. The paradox is that current AI narratives denote a failure of imagination by making use of age-old international relations tropes and concepts of great power competition, geopolitics, arms races, and existential threats on a par with nuclear annihilation or pandemics. The EU has also fallen into these narrative traps; whether true or false, they engender self-fulfilling prophecies and paint possible futures as inevitable, thus limiting decisionmakers' options to respond.

As AI systems evolve, the EU must grapple with questions of transparency, responsibility, and ethical frameworks to ensure that the union's technopolitical approaches align with broader ethical, democratic, and socially just goals. Starting with the launch of ChatGPT in late 2022, the EU governance landscape witnessed a whirlwind year that culminated in a significant agreement on the AI Act in December 2023. While the heightened momentum and discourses on AI governance are encouraging and important, the pivotal question for 2024 is whether these discussions will yield tangible global and European commitments when it comes to governing AI, addressing paramount AI risks, and, crucially, translating goals into substantive action across jurisdictions. Given that AI developments present new opportunities and benefits, the consequences of a perceived geopolitical race to innovate and regulate AI are potentially more far-reaching than the question of who wins global hegemony, whether in the case of states or corporate players.

The perception of an AI arms race is likely to accelerate the already risky development of AI systems. The pressure to outpace public and private sector adversaries by rapidly pushing the frontiers of a technology that is still not fully understood or controlled, and without commensurate efforts to make AI safe for humans, may well present insurmountable challenges and risks. A détente in this race, however improbable it may seem today, may be crucial to humanity's long-term prosperity and safety.

In this respect, the EU faces a big test to manage the geopolitics of AI governance in a landscape characterized by state and corporate competition as well as an emerging global regime complex of regulatory frameworks. While the EU's commitment to responsible AI is commendable, building a harmonized EU foreign policy approach to AI, fostering strategic alliances with key partners, operationalizing the AI Act effectively, and navigating diverse governance initiatives will be crucial for shaping the future of AI on a global scale. Importantly, the EU will need to evaluate how well its AI governance lives up to its normative ideals and European values.

About the Author

Raluca Csernatoni is fellow at Carnegie Europe, where she specializes on European security and defense, as well as emerging disruptive technologies. She is also a team leader and research expert on new technologies for the EU Cyber Direct - EU Cyber Diplomacy Initiative project.

Acknowledgments

The author would like to thank Rosa Balfour, Matt Sheehan, Hadrien Pouget, and Chantal Lavallée for providing insightful feedback on earlier drafts of this working paper.

Carnegie Europe is grateful to the Patrick J. McGovern Foundation for their support of this work.

Notes

- 1 Henry Farrell and Abraham L. Newman, "Weaponized Interdependence: How Global Economic Networks Shape State Coercion," *International Security* 44, no. 1 (2019): 42–79, https://doi.org/10.1162/isec_a_00351.
- 2 Matt O'Shaughnessy, "How Hype Over AI Superintelligence Could Lead Policy Astray," Carnegie Endowment for International Peace, September 14, 2023, https://carnegieendowment.org/2023/09/14/how-hype-over-ai-superintelligence-could-lead-policy-astray-pub-90564.
- Jascha Bareis and Christian Katzenbach, "Talking AI Into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics," *Science, Technology, & Human Values* 47, no. 5 (2022): 855–881, https://doi.org/10.1177/01622439211030007.
- 4 Barry Pavel et al., "AI and Geopolitics: How Might AI Affect the Rise and Fall of Nations?," RAND Corporation, November 3, 2023, https://www.rand.org/pubs/perspectives/PEA3034-1.html.
- 5 "A European Approach to Artificial Intelligence," European Commission, January 31, 2024, https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.
- 6 "Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence," European Commission, April 21, 2021, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence.
- 7 Ian Manners, "Normative Power Europe: A Contradiction in Terms?," *Journal of Common Market Studies* 40, no. 2 (2002): 235–258, https://doi.org/10.1111/1468-5965.00353.
- 8 "Disruptive Technology Calls," European Union, https://eudis.europa.eu/disruptive-technology-calls en.
- 9 Vera Bergengruen, "How Tech Giants Turned Ukraine Into an AI War Lab," *Time*, February 8, 2024, https://time.com/6691662/ai-ukraine-war-palantir/.
- 10 Raluca Csernatoni, "The EU's Technological Power: Harnessing Future and Emerging Technologies for European Security," in *Peace, Security and Defence Cooperation in Post-Brexit Europe: Risks and Opportunities*, ed. Cornelia-Adriana Baciu and John Doyle (New York: Springer, 2019), 119–140.
- Nestor Maslej et al., "Artificial Intelligence Index Report 2023," Stanford University, April 2023, 6, https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

- 12 "Artificial Intelligence, Blockchain and the Future of Europe," European Investment Bank, June 1, 2021, https://www.eib.org/en/publications/artificial-intelligence-blockchain-and-the-future-of-europe-report.
- "Action Plan on Synergies Between Civil, Defence and Space Industries," European Commission, February 22, 2021, 13, https://commission.europa.eu/system/files/2021-03/action_plan_on_synergies_en_1.pdf.
- Definition added. See Charmaine Chua, "Disruption From Above, the Middle and Below: Three Terrains of Governance," Review of International Studies 49, no. 1 (2023): 37-52, https://doi.org/10.1017/ S0260210522000432.
- Raluca Csernatoni and Bruno Oliveira Martins, "Disruptive Technologies for Security and Defence: Temporality, Performativity and Imagination," Geopolitics (2023), https://doi.org/10.1080/14650045.2023.
- 16 Brian Mullins, "AI, Super Intelligence, and the Fear of Machines in Control," The Cyber Defense Review 7, no. 2 (2022): 67-76, https://www.jstor.org/stable/48669293.
- "2023 State of the Union Address by President Von der Leyen," European Commission, September 13, 2023, https://ec.europa.eu/commission/presscorner/detail/ov/speech 23 4426.
- Raluca Csernatoni and Katerina Mavrona, "The Artificial Intelligence and Cybersecurity Nexus: Taking Stock of the European Union's Approach," Carnegie Europe and EU Cyber Direct, September 15, 2022, https://carnegieeurope.eu/2022/09/15/artificial-intelligence-and-cybersecurity-nexus-taking-stock-of-european-union-s-approach-pub-87886.
- Matt Sheehan, "China's AI Regulations and How They Get Made," Carnegie Endowment for International Peace, July 10, 2023, https://carnegieendowment.org/2023/07/10/china-s-ai-regulations-and-howthey-get-made-pub-90117.
- Raluca Csernatoni, "Beyond the Hype: The EU and the AI Global 'Arms Race," European Leadership Network, August 21, 2019, https://carnegieeurope.eu/2019/08/21/beyond-hype-eu-and-ai-globalarms-race-pub-79734.
- Josh Luckenbaugh, "New Pentagon Task Force Exploring Generative AI," National Defense, October 25, 2023, https://www.nationaldefensemagazine.org/articles/2023/10/25/new-pentagon-task-forceexploring-generative-ai.
- Maslej et al., "Artificial Intelligence Index Report 2023."
- 23 Anu Bradford and Raluca Csernatoni, "Toward a Strengthened Transatlantic Alliance," in Working With the Biden Administration: Opportunities for the EU, ed. Rosa Balfour, Carnegie Endowment for International Peace, January 26, 2021, https://carnegieendowment.org/2021/01/26/ toward-strengthened-transatlantic-technology-alliance-pub-83565.
- Ulrike Franke, "Artificial Intelligence Diplomacy: Artificial Intelligence Governance as a New European Union External Policy Tool," European Parliament, June 2021, https://www.europarl.europa.eu/RegData/ etudes/STUD/2021/662926/IPOL STU(2021)662926 EN.pdf.
- Luca Bertuzzi, "Mastery of Technology Is Central to the 'New Geopolitical Order', Breton Says," Euractiv, July 27, 2021, https://www.euractiv.com/section/industrial-strategy/news/ mastery-of-technology-is-central-to-the-new-geopolitical-order-breton-says.
- "Speech by President-Elect Von der Leyen in the European Parliament Plenary on the Occasion of the Presentation of her College of Commissioners and Their Programme," European Commission, November 27, 2019, https://ec.europa.eu/commission/presscorner/detail/en/SPEECH 19 6408.
- Franke, "Artificial Intelligence Diplomacy."
- "Welcome to the Era of AI Nationalism," Economist, January 1, 2024, https://www.economist.com/ business/2024/01/01/welcome-to-the-era-of-ai-nationalism.
- Maslej et al., "Artificial Intelligence Index Report 2023."
- Raluca Csernatoni, "Towards Strengthening the Transatlantic Tech Diplomacy: Trustworthy AI in the EU-U.S. Trade and Technology Council," Transatlantic Leadership Network, January 2023, https://www. transatlantic.org/wp-content/uploads/2023/01/Csernatoni Background-Paper-on-the-EU-US-TTC-transatlantic.org/wp-content/uploads/2023/01/Csernatoni Background Cooperation-on-AI.pdf?mc_cid=5c3d87eca1.

- 31 Jamil Anderlini and Clea Caulcutt, "Europe Must Resist Pressure to Become 'America's Followers,' Says Macron," Politico, April 9, 2023, https://www.politico.eu/article/emmanuel-macron-china-americapressure-interview.
- 32 "EU-US Trade and Technology Council," European Commission, https://commission.europa.eu/ strategy-and-policy/priorities-2019-2024/stronger-europe-world/eu-us-trade-and-technology-council_en.
- "Commission Welcomes G7 Leaders' Agreement on Guiding Principles and a Code of Conduct on Artificial Intelligence," European Commission, October 30, 2023, https://ec.europa.eu/commission/presscorner/ detail/en/ip 23 5379.
- Nathalie A. Smuha, "From a 'Race to AI' to a 'Race to AI Regulation' Regulatory Competition for Artificial Intelligence," Law, Innovation and Technology 13, no. 1 (2021): 57-84, https://dx.doi.org/10.2139/ ssrn.3501410.
- Anu Bradford, The Brussels Effect: How the European Union Rules the World (Oxford: Oxford University Press,
- Matthijs M. Maas and José Jaime Villalobos, "International AI Institutions: A Literature Review of Models, Examples, and Proposals," Legal Priorities Project, September 23, 2023, 7, https://dx.doi.org/10.2139/ ssrn.4579773.
- "OECD AI Principles Overview," Organisation for Economic Co-operation and Development, https:// oecd.ai/en/ai-principles; "About GPAI," Global Partnership on Artificial Intelligence, https://gpai.ai/ about; "Recommendation on the Ethics of Artificial Intelligence," United Nations Educational, Scientific, and Cultural Organization, 2021, https://unesdoc.unesco.org/ark:/48223/pf0000380455; and "Ethics Guidelines for Trustworthy AI," European Commission, April 8, 2019, https://digital-strategy.ec.europa.eu/ en/library/ethics-guidelines-trustworthy-ai.
- "Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law," Council of Europe, December 18, 2023, https://rm.coe.int/cai-2023-28-draft-framework-convention/ 1680ade043.
- "AI Governance Alliance," World Economic Forum, https://initiatives.weforum.org/ai-governance-alliance/
- 40 "Tech Accord to Combat Deceptive Use of AI in 2024 Elections," Munich Security Conference, https:// securityconference.org/en/aielectionsaccord.
- 41 Maslej et al., "Artificial Intelligence Index Report 2023."
- 42 "AI Risk Management Framework," U.S. National Institute of Standards and Technology, https://www.nist. gov/itl/ai-risk-management-framework; and "Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People," White House, October 2022, https://www.whitehouse.gov/wp-content/ uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.
- "Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence," White House, October 30, 2023, https://www.whitehouse.gov/briefing-room/presidentialactions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-ofartificial-intelligence.
- Sheehan, "China's AI Regulations."
- "G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI," Organisation for Economic Co-operation and Development, September 7, 2023, https://www.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative -artificial-intelligence-ai_bf3c0c60-en.
- Mariano-Florentino (Tino) Cuéllar, "The UK AI Safety Summit Opened a New Chapter in AI Diplomacy," Carnegie Endowment for International Peace, November 9, 2023, https://carnegieendowment. org/2023/11/09/uk-ai-safety-summit-opened-new-chapter-in-ai-diplomacy-pub-90968.
- "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023," United Kingdom Government, November 1, 2023, https://www.gov.uk/government/publications/ ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attendingthe-ai-safety-summit-1-2-november-2023.

- "Secretary-General Announces Creation of New Artificial Intelligence Advisory Board," United Nations, October 26, 2023, https://press.un.org/en/2023/sga2236.doc.htm.
- "Interim Report: Governing AI for Humanity," United Nations AI Advisory Body, December 2023, https:// www.un.org/sites/un2.un.org/files/ai advisory body interim report.pdf.
- Luca Bertuzzi, "EU Prepares to Push Back on Private Sector Carve-Out From International AI Treaty," Euractiv, January 10, 2024, https://www.euractiv.com/section/artificial-intelligence/news/ eu-prepares-to-push-back-on-private-sector-carve-out-from-international-ai-treaty.
- Shazeda Ahmed et al., "Building the Epistemic Community of AI Safety," November 22, 2023, https://dx.doi.org/10.2139/ssrn.4641526.
- "High-Level Expert Group on Artificial Intelligence," European Commission, June 7, 2022, https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.
- "Communication [on] Artificial Intelligence for Europe," European Commission, April 25, 2018, https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe; and "Ethics Guidelines," European Commission.
- "Policy and Investment Recommendations for Trustworthy Artificial Intelligence," European Commission, June 26, 2019, https://digital-strategy.ec.europa.eu/en/library/policy-and-investmentrecommendations-trustworthy-artificial-intelligence.
- "White Paper on Artificial Intelligence: A European Approach to Excellence and Trust," European Commission, February 19, 2020, https://commission.europa.eu/publications/white-paper-artificialintelligence-european-approach-excellence-and-trust en.
- "About," European Commission, https://ai-watch.ec.europa.eu/about_en.
- "AI Watch: Estimating AI Investments in the European Union," European Commission, May 23, 2022, 3, https://ai-watch.ec.europa.eu/publications/ai-watch-estimating-ai-investments-european-union_en.
- "Communication From the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial Intelligence for Europe," European Commission, June 26, 2018, https://ec.europa.eu/transparency/documents-register/ detail?ref=COM(2018)237&lang=en.
- "AI Watch," European Commission.
- "Coordinated Plan on Artificial Intelligence 2021 Review," European Commission, April 21, 2021, https:// digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review.
- "2022 Strategic Foresight Report 'Twinning the Green and Digital Transitions in the New Geopolitical Context," European Commission, July 1, 2022, https://knowledge4policy.ec.europa.eu/publication/2022-strategic-foresight-report-%E2%80%9Ctwinning-green-digital-transitions-new-geopolitical en.
- "Security Europe's Competitiveness: Addressing Its Technology Gap," McKinsey Global Institute, September 22, 2022, 10, https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/ securing-europes-competitiveness-addressing-its-technology-gap.
- "Artificial Intelligence," European Investment Bank.
- "Joint Communication to the European Parliament, the European Council and the Council on 'European Economic Security Strategy," European Commission, June 20, 2023, https://eur-lex.europa.eu/ legal-content/EN/TXT/?uri=CELEX%3A52023JC0020&qid=1687525961309.
- "Joint Non-paper on Open Strategic Autonomy of the EU, Signed by Belgium, Finland, Portugal, Slovakia and the Netherlands," Permanent Representation of the Kingdom of the Netherlands to the European Union, July 19, 2023, https://open.overheid.nl/documenten/5f3a6437-92b3-41bc-835a-e4d803ee6f6b/file.
- "Commission Recommendation of 03 October 2023 on Critical Technology Areas for the EU's Economic Security for Further Risk Assessment With Member States," European Commission, October 3, 2023, https://defence-industry-space.ec.europa.eu/commission-recommendation-03october-2023-critical-technology-areas-eus-economic-security-further en.

- "Communication on Boosting Startups and Innovation in Trustworthy Artificial Intelligence," European Commission, January 24, 2024, https://digital-strategy.ec.europa.eu/en/library/ communication-boosting-startups-and-innovation-trustworthy-artificial-intelligence.
- "2023 State of the Union Address," European Commission.
- "Communication on Boosting Startups," European Commission, 4.
- Sébastien Bubeck et al., "Sparks of Artificial General Intelligence: Early Experiments With GPT-4," March 22, 2023, ArXiv, https://doi.org/10.48550/arXiv.2303.12712.
- Kevin Roose, "A.I. Poses 'Risk of Extinction,' Industry Leaders Warn," New York Times, May 30, 2023, https://www.nytimes.com/2023/05/30/technology/ai-threat-warning.html.
- "AI Act," European Commission, February 20, 2024, https://digital-strategy.ec.europa.eu/en/policies/ regulatory-framework-ai.
- "Artificial Intelligence Act: Deal on Comprehensive Rules for Trustworthy AI," European Parliament, December 9, 2023, https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/ artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai.
- "Commission Welcomes Political Agreement on Artificial Intelligence Act," European Commission, December 9, 2023, https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473.
- "Commission Decision Establishing the European AI Office," European Commission, January 24, 2024, https://digital-strategy.ec.europa.eu/en/library/commission-decision-establishing-european-ai-office.
- "Note From the Presidency to the Permanent Representatives Committee," Council of the European Union, January 26, 2024, https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf.
- Gian Volpicelli, "EU Set to Allow Draconian Use of Facial Recognition Tech, Say Lawmakers," Politico, January 16, 2024, https://www.politico.eu/article/eu-ai-facial-recognition-tech-act-latetweaks-attack-civil-rights-key-lawmaker-hahn-warns.
- "Prohibit Remote Biometric Categorisation in Publicly Accessible Spaces, and Any Discriminatory Biometric Categorisation," Access Now, November 2021, https://www.accessnow.org/wp-content/uploads/2022/05/ Amendments-to-the-AI-Acts-treatment-of-biometric-categorisation.pdf.
- "EU AI Act Will Fail Commitment to Ban Biometric Mass Surveillance," Reclaim Your Face, January 18, 2024, https://reclaimyourface.eu/eu-ai-act-will-fail-commitment-to-ban-biometric-mass-surveillance.
- Alexandre Piquard, "France Keeps Up Pressure on EU's AI Act, Despite Mounting Criticism," Le Monde, January 27, 2024, https://www.lemonde.fr/en/economy/article/2024/01/27/france-keeps-up-its-pressure-onthe-eu-s-ai-act-despite-mounting-criticism_6471038_19.html.
- Julia Tar and Luca Bertuzzi, "AI Office Established, AI Convention's Scope Struggle," Euractiv, January 26, 2024, https://www.euractiv.com/section/digital/news/ai-office-established-ai-conventions-scope-struggle.
- "Artificial Intelligence Act: Committees Confirm Landmark Agreement," European Parliament, February 13, 2024, https://www.europarl.europa.eu/news/en/press-room/20240212IPR17618/ artificial-intelligence-act-committees-confirm-landmark-agreement.
- "Artificial Intelligence Act," European Parliament.

Carnegie Europe

Carnegie Europe delivers interdisciplinary expertise and independent insights that bring together national, regional, and global perspectives and help European policymakers grasp and respond to global challenges. From Brussels, we focus on three themes: the European Union's relations with its partners and competitors; the risks of democratic backsliding in Europe and around the world; and Europe's efforts to meet today's most pressing global challenges, from climate change to the new frontiers of cyber diplomacy.

Carnegie Endowment for International Peace

The Carnegie Endowment for International Peace is a unique global network of policy research centers around the world. Our mission, dating back more than a century, is to advance peace through analysis and development of fresh policy ideas and direct engagement and collaboration with decisionmakers in government, business, and civil society. Working together, our centers bring the inestimable benefit of multiple national viewpoints to bilateral, regional, and global issues.

