

3

The policies of AI and trade

(a) AI and trade: key policy considerations

The discussion of how AI might reshape international trade raises important policy questions. The future of AI and international trade hinges on the policy choices of governments and on the strategies and priorities of industries and businesses.

(i) Addressing the growing AI divide

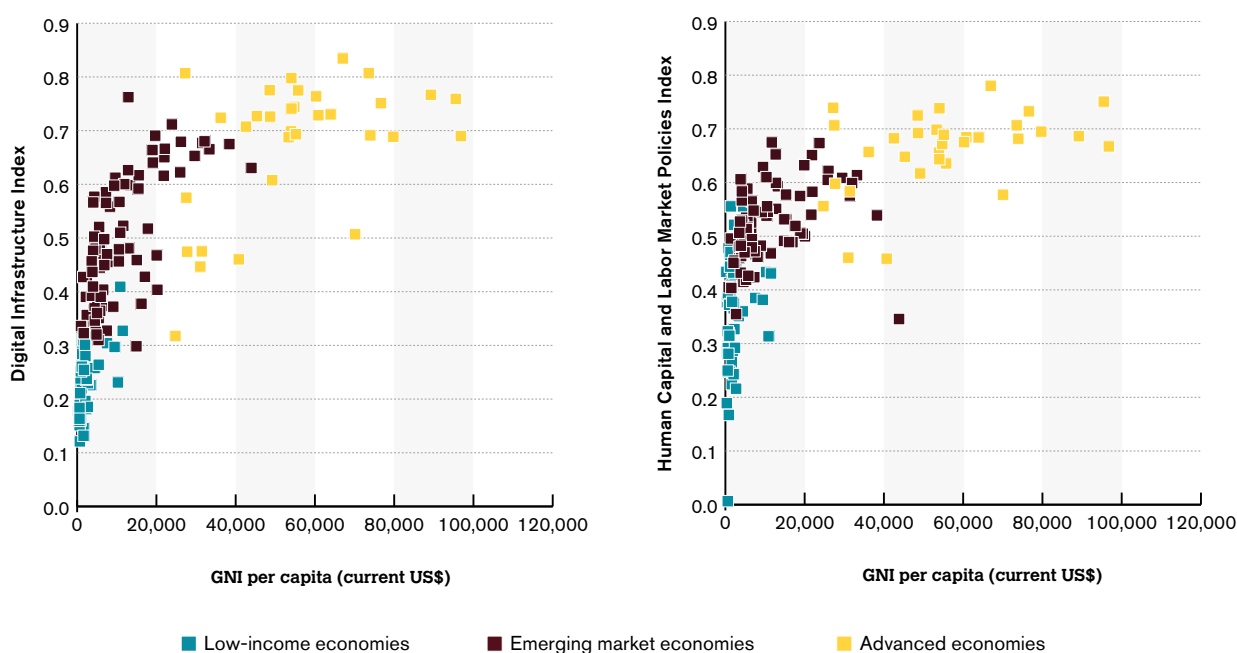
To leverage the opportunities of AI, the digital divide between economies, in terms of both digital infrastructure and skills, must be addressed. As discussed in Chapter 2, ensuring that workers and firms are prepared to adopt AI involves robust digital infrastructure and trained human capital. Digital infrastructure is a crucial determinant of information and communications technology (ICT) adoption, and can lay the foundation for the diffusion and localized application of AI technology (Nicoletti et al., 2020). Nonetheless, such infrastructure is of limited use without a skilled workforce capable of leveraging digital platforms for innovative workplace applications.

Wealthier economies, including advanced and some emerging market economies, are generally better prepared than low-income economies to adopt AI.

As illustrated in Figure 3.1, both the Digital Infrastructure Index and the Human Capital and Labor Market Policies Index — components of the IMF's AI Preparedness Index — are positively correlated with income levels. Higher-income economies tend to have stronger digital infrastructure and more trained human capital, making them more equipped to adopt AI technologies.

To address the AI divide, it is crucial to invest in digital infrastructure to ensure that low-income economies have the necessary technological foundation to support AI adoption. Governments and the private sector could collaborate to expand high-speed internet access, improve electricity infrastructure, particularly through renewable energy generation, enhance data storage capabilities, and develop robust cybersecurity measures. Public policies need to incentivize infrastructure development in underserved areas, and international cooperation should focus on providing technical and financial assistance to developing economies. In addition to infrastructure, bridging the AI divide requires a substantial investment in human capital to equip individuals with the skills they need to utilize AI technologies effectively. Education and training programmes should include AI literacy, coding, data analysis and other relevant skills. Public-private partnerships can play a key role in these efforts, as companies can offer practical training and resources, while governments can provide the necessary regulatory support, access to affordable devices and connectivity, and funding (see the opinion piece by James Manyika).

Figure 3.1: AI preparedness is higher in advanced economies



Source: International Monetary Fund (IMF) AI Preparedness Index (Cazzaniga et al., 2024). The indices are rescaled to range between 0 and 1.

Opinion piece

Harnessing technology to advance shared prosperity

Artificial intelligence is the most important technology of the present era, offering the potential to make people's everyday lives easier, power economic growth, help middle-class and lower-income workers, drive scientific and health advances, and address longstanding development challenges. In a world that is on track to meet only 15 per cent of the UN's Sustainable Development Goals (SDGs), AI provides an opportunity to reverse that trendline and contribute to progress on 79 per cent of our shared global goals (Hoyer Gosselink et al., 2024).

However, while the economic and societal opportunity offered by AI is immense, we must always remember that the benefits of new technologies are not automatic. At this moment of excitement, it is important to take a step back and consider the history of trade and technology – and make a concerted effort to build an inclusive trading system around AI that avoids the creation of an “AI divide.”

The combined forces of technology and trade helped lift over a billion people out of extreme poverty – an achievement unparalleled in human history.¹ This dramatically changed global development, speeding up the flow of data and enabling smaller companies and economies to participate in trade. By 2016, digital flows – often a key aspect of other global flows such as manufacturing, services and financial flows and other intangibles – had begun to exert a larger impact on GDP growth than the centuries-old trade in goods (Manyika et al., 2016), with a surprising share of the benefits from digital trade going to the services, manufacturing and retail sectors (The White House, 2024). These trends on digital and digitally enabled global flows have only accelerated since 2016, with some estimating that up to 40 per cent of GDP now depends on global flows (Seong et al., 2022).

Now with AI, economies stand on the verge of an even more profound economic and scientific transformation that may fundamentally shift and reshape trading patterns. But it is critical to avoid creating an “AI divide” (Ossa, 2023). As of 2023, 93 per cent of people in high-income economies use the internet, compared with only 27 per cent of people in low-income economies (ITU, 2023). The UN's AI Advisory Body has rightly concluded that these divides cannot persist into the AI era (United Nations, 2024a).

We must work together across companies, governments and civil society to harness AI to advance a vision of shared prosperity. At a national level,

James Manyika

Senior Vice President of Research,
Technology and Society at Google;
Co-Chair of the United Nations
Secretary-General's High-Level
Advisory Body on
Artificial Intelligence

this means providing access to AI-capable cloud infrastructure, computing capacity, developer tools and datasets relevant to AI development, while equipping workers and students with foundational AI skills that provide pathways to the modern workforce. Most crucially, we must place small businesses and traditional industries like manufacturing and agriculture at the forefront of AI leadership.

At the international level, driving a vision of shared prosperity will require expanding what we think of as “trade” – not just removing barriers to cross-border goods and services, but advancing a global strategy to drive alignment on AI governance and security, support trusted data flows, enable economic integration, and build solutions to cross-border challenges. An inclusive AI strategy must also drive investment in the subsea and terrestrial cables that enable participation in the modern economy (Quigley, 2024).

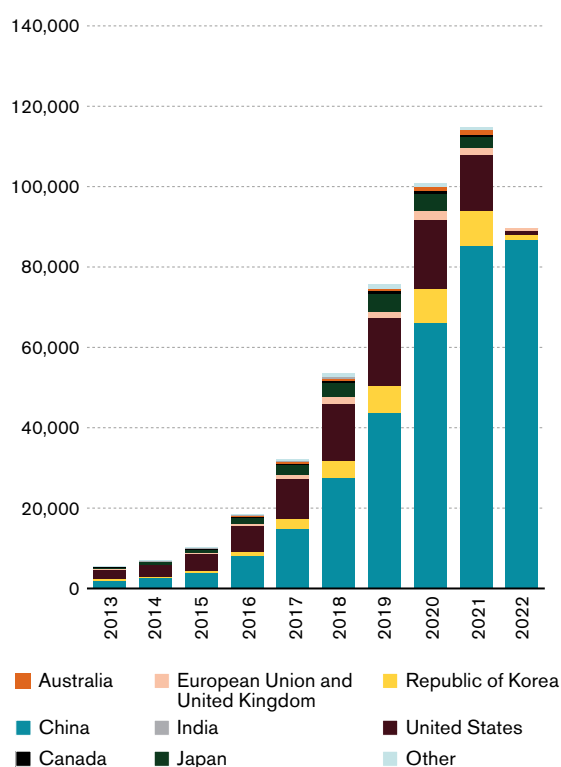
At its heart, this is a modern form of capacity building, with governments, industry and civil society working together to invest in AI infrastructure, build a global resource for AI research and develop training programmes that promote AI diffusion across sectors and geographies.

In contrast, if economies cannot align trade with the mission of shared prosperity, there is a risk that AI will only be adopted by wealthier economies, and by the wealthiest industries within those economies. This would be harmful not just from an equity perspective but also from an economic perspective – the trillions in potential economic benefits from AI are conditional on broad-based adoption of these technologies, not usage by the privileged few.

The choice is ours. Together, let's build a trade agenda that harnesses the transformative power of AI for all people, regardless of geographical location or economic status.

Disclaimer

Opinion pieces are the sole responsibility of their authors. They do not necessarily reflect the opinions or views of WTO members or the WTO Secretariat.

Figure 3.2: Number of AI patent filings by geographic regions, 2013-22

Source: Centre for Security and Emerging Technology (2024).

In addition to differences in the ability to adopt AI technologies, the AI divide across economies, reflected in AI research and development (R&D), investment and expertise, highlights the need to address gaps in AI capabilities. There is a significant divide between economies leading in AI R&D and the rest, especially developing economies and least-developed countries (LDCs).

As illustrated in Figure 3.2, China is by far the leading economy in terms of the number of patents registered, with 86,663 AI patent applications and 15,869 patents granted in 2022, followed by the Republic of Korea and the United States.² This disparity reflects the underlying technological differences between economies and underscores the importance of facilitating technology dissemination and technical assistance to bridge the gap globally. There is a marked division between where the research, patents and investments in AI are located and where they are lacking, and there is a growing risk of further exacerbating this division, which exists both between and within economies and between urban and rural, less digitally connected areas.

The number of published articles on AI has increased steadily, with industry taking the lead. The number of published articles on AI has increased steadily, except in the United States which saw a drop in 2022. Although China is leading in terms of the volume of published scholarly articles, it is worth noting that the United States ranks first in terms of the number of citations, an indicator of the

literature's influence.³ As illustrated in Figure 3.3, the affiliation of teams researching and publishing on AI systems that “demonstrate the ability to learn, show tangible experimental results, and contribute advancements that push the boundaries of existing AI technology” showed an important switch in the second part of the 2010s: whereas most of the research was led by academia prior to 2016, industry took the lead in the number of publications afterwards.

Investment in AI is accelerating rapidly, with the United States leading in private investment. AI funding stems from a variety of sources, including private companies, venture capital firms, government funding, academic institutions, corporate partnerships and angel investors. The United States leads in terms of total AI private investment.

In 2023, the US\$ 67.2 billion invested in AI in the United States was roughly 8.7 times greater than the amount invested in the next highest country, China (US\$ 7.8 billion), and 17.8 times the amount invested in the United Kingdom (US\$ 3.8 billion). Since 2022, the United States has experienced a notable increase in private investment in AI (22.1 per cent) (Maslej et al., 2024).

The disparity in private investment in AI becomes particularly pronounced in generative AI. Despite a recent decline in overall AI private investment, funding for generative AI has surged, reaching US\$ 25.2 billion in 2023 (Maslej et al., 2024).

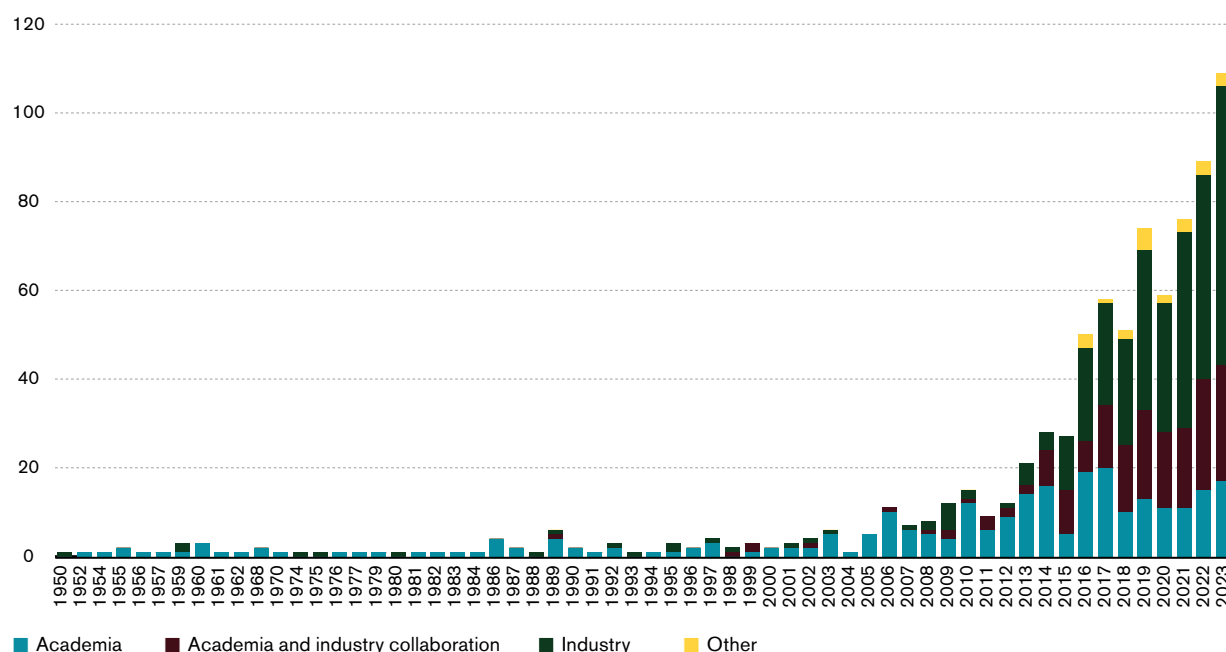
However, this surge is heavily concentrated in a few economies, with the United States taking the lead. In 2023, the United States surpassed the combined investments of the European Union plus the United Kingdom in generative AI by approximately US\$ 21.1 billion. Venture capital investments in generative AI have also been led by the United States, with a steep jump in 2023 to over US\$ 16 billion going towards generative adversarial networks (machine learning models that generate new data mimicking a given dataset) for AI training and generative AI for text, image and audio.⁴

The demographics of professionals with AI skills are largely male and located in Europe and North America.

According to a developer survey by Stack Overflow, a question-and-answer platform for programmers,⁵ 94.24 per cent of data scientists and machine learning professionals are male, and the majority are located in Europe and North America (OECD.AI, 2024). Moreover, data from the Computing Research Association,⁶ although limited to the United States and Canada, reveal that the representation of women among new AI and computer science PhDs has remained stagnant at approximately 20 per cent since 2010. This persistent gender gap underscores an ongoing challenge within the field.

This imbalance may be further exacerbated by the race to nurture AI through government subsidies.

As discussed in Chapter 3(b), a number of governments are launching domestic initiatives to promote AI, backed by generous state support. However, as most of this support is being provided by high-income economies, it may exacerbate the AI divide among economies. The relative concentration of AI supply chains can also result in trade imbalances in AI-related goods and services.

Figure 3.3: Affiliation of publication research teams, 1950-2022

Source: Epoch (2024).

Beyond the digital divide across economies, industrial concentration is prevalent in AI within economies due to increasing returns and network effects.

As discussed in Section 2(b)(iv), as the development of AI models progresses and their development costs escalate, only large firms can afford the substantial up-front investments required. This creates a significant barrier for newcomers and smaller enterprises, making it difficult for them to compete. The high initial costs of developing cutting-edge AI models and the necessity for extensive data and computational resources further consolidate the dominance of established players.

The widespread adoption of AI in markets can heighten the risk of collusion between companies.

AI systems integrated into pricing strategies and market analysis can enable companies to monitor competitors' pricing behaviour and adjust their own prices accordingly. While this may optimize profits individually, it can collectively lead to tacit agreements or collusion among competitors to maintain higher prices (Assad et al., 2024; OECD, 2021a). Moreover, AI's ability to process vast data and predict market trends may enhance firms' coordination in pricing strategies, exacerbating market concentration.⁷

The special features of AI present challenges for competition authorities.

The opacity of AI algorithms and the sheer volume of data they process can obscure anti-competitive practices such as price collusion, exclusionary behaviour and discriminatory practices. Moreover, AI-driven mergers and acquisitions may raise concerns about market dominance and barriers to entry, as algorithms and data assets become pivotal assets for competitive

advantage. Traditional antitrust frameworks may struggle to adapt to the dynamic nature of AI-driven markets, requiring competition authorities to develop new analytical tools, data access mechanisms and regulatory frameworks to effectively safeguard competition and consumer welfare in the AI era (see the opinion piece by Shin-yi Peng).

There is growing scrutiny of mergers in the AI market and growing interest in better understanding the implications of AI on competition.⁸

Traditional antitrust policies, which apply after the fact, when market competition has already been impacted, are slow and focus on prices, and are not sufficient to address competition issues raised by AI. The competition challenges raised by AI have led to renewed calls for a collective international approach to regulation and for enforcement of competition in digital markets.

(ii) Preventing further digital trade barriers

Cross-border data flows are essential to AI.

As discussed in Section 2(a)(i), amassing vast datasets is vital in order to train algorithms, and data flows are integral to the real-time use of AI technologies. Breadth and variety of data are as important as volume.⁹ For AI to be effective and deliver accurate predictions that are not susceptible to bias and discrimination, algorithms need to be built on high-quality, accurate and representative data.

Cross-border data flow restrictions can negatively impact AI innovation and development, and can increase costs for firms.

Such restrictions have a general negative impact on productivity, economic growth and innovation domestically and globally (Aaronson, 2019; Goldfarb and Tucker, 2012; Luintel and Khan, 2009; Maskus and Reichman, 2004; OECD, 2016), but are of particular concern for AI innovation and development. Because AI requires vast amounts of good quality data in order to be trained, and this often involves merging different data sources together, cross-border data flow restrictions are likely to affect the quality and accuracy of AI models and the scalability of AI applications significantly. By limiting the ability of foreign firms to access data from a given jurisdiction, such measures could favour domestic

firms, but may do so at the expense of overall quality, thereby undermining innovation and the full potential of AI (Goldfarb and Trefler, 2018).¹⁰ Cross-border data flow restrictions also impose extra costs on firms wanting to do business internationally. A recent study on the implications of data flow restrictions on global GDP and trade finds that if all economies fully restricted their data flows, it could result in a 5 per cent reduction in global GDP and a 10 per cent decrease in exports (OECD and WTO, 2024). To comply with data flow restrictions, firms may need to establish a presence and duplicate activities across various jurisdictions and devise a system to ensure that data are not routed internationally. While technically feasible, doing this can be particularly costly, especially for small businesses (Goldfarb and Trefler, 2018).

Box 3.1: AI and consumer protection

AI opens significant opportunities for consumers, but also increases the possibilities of covert influence, raising significant concerns over the exploitation of personal information and violation of privacy, manipulation and disinformation.

For consumers, AI can provide major benefits, such as individualized recommendations and time-saving (e.g., AI voice assistants can order groceries instantly, saving consumers hours of shopping time). The ability of AI models to establish correlations between consumers' data and possible responses to advertisements in order to predict consumers' behaviour provides firms using AI with the unprecedented ability to trigger specific reactions through individualized advertisements and communications. However, this can also exacerbate asymmetry of information between companies and consumers, and can lead to manipulation and exploitation of consumer behaviour.

The use of algorithms to fix prices can lead to price efficiencies passed on to customers, but can also be used to exploit consumers' willingness to pay a certain price in the interest of the firm.

In addition, while AI technologies can contribute to effective moderation for the benefit of consumers, they can also deliver inaccurate, biased or discriminatory responses that can also harm consumers.

Finally, sellers may not fully take into account potential harm caused to consumers as a result of consumer data misuse due to the difficulty in tracing that harm back to the original data collector. Consumers may not, therefore, challenge data use after the data is collected (Agrawal et al., 2019).

Notwithstanding the fact that traditional consumer protections laws may apply to most scenarios of AI use cases, often providing adequate legal remedies without the need for new regulations, measures to regulate algorithmic harm and protect consumers have emerged in recent years in various jurisdictions. Under the EU General Data Protection Regulation, for example, individuals have the right to contest decisions made by algorithms, request human oversight, and withdraw from personalized advertising driven by algorithmic methods. China has also developed comprehensive regulations to govern algorithm use. Further legislation

to harness algorithms and protect consumers is being discussed in various jurisdictions, including the European Union¹² and the United Kingdom (Holmes, 2024).

However, national approaches do not adequately protect consumers in the case of cross-border transactions (Jones, 2023). Obtaining redress in case of harm remains particularly challenging in the event of international transactions. Although some level of international collaboration and regulatory discussions exists among national bodies, this cooperation remains fragmented and does not establish an effective, transparent framework for enforcing consumer rights across borders (Goyens, 2020), leading some experts to call for new forms of international regulation and cooperation to protect consumers, especially against AI harm (Jones, 2023).

Besides potential violations of privacy and personal integrity, disinformation and manipulation, the difficulty of assessing the safety and security of AI-enabled products and services adds to the complexity of protecting consumers in an AI-driven age (see Chapter 3(a)(iii)).

Opinion piece

AI: Amplifying the digital trade issues

Many of the challenges brought by digital technologies that the WTO has faced in the past decades are now amplified by AI, including issues associated with classification, non-discrimination, data governance and competition policy.

First of all, AI acts as a facilitator for complex products that bundle goods and services, which calls for further thinking about how to adjust the goods/services legal silos under the WTO to address issues stemming from the merging of physical and digital realms. It seems likely that the goods/services dichotomy in applying trade rules will increasingly trigger new levels of inconsistency and legal uncertainty.

Second, more and more AI-based services will be able to compete directly with or substitute human professionals. Questions such as to what extent automated legal advice tools and human attorneys should be considered to be “like services suppliers” may emerge sooner than expected. It remains to be seen how far concepts such as “technological neutrality” or “evolutionary interpretation” can serve to clarify the scope of the GATS commitments of market access and national treatment.

Third, AI presents new challenges for data governance. Digital platforms’ advertising algorithms, or, more generally, their overall business models, intensify the perils associated with the data-driven economy. WTO rules can play a more active role in reducing such perils.

Shin-yi Peng

Distinguished Professor of Law,
National Tsing Hua University

Data governance in the age of AI requires perspectives that safeguard social values, including privacy, security, free speech, cultural expression and algorithmic ethics.

Finally, competition authorities worldwide are increasingly taking or considering approaches that impose additional obligations on AI-powered big tech companies. Their potential abuses of market power, including self-preferencing practices that promote their own services within search results, algorithmic cartels and other cross-border collusive arrangements, can be more meaningfully addressed through competition disciplines at the international level. To what extent does algorithmic practice constitute a trade barrier to goods or services, and how do anti-competitive market concentrations exclude foreign suppliers from a market? The reactivation of the WTO Working Group on Trade and Competition is more urgent than ever. If a set of general or sector-specific competition disciplines could be established at the WTO, it would be less necessary for competition authorities in developing countries and LDCs to enforce competition law after the fact.

Disclaimer

Opinion pieces are the sole responsibility of their authors. They do not necessarily reflect the opinions or views of WTO members or the WTO Secretariat.

However, the large datasets required by AI models raise significant privacy concerns. AI introduces new privacy issues for individuals and consumers, leading to a trade-off between the necessity of accessing large amounts of data to train AI models and privacy protection. The continuous tracking and profiling of individuals’ online and offline interactions by AI algorithms raise significant concerns about data privacy, consent and control over personal information. Furthermore, as AI algorithms become increasingly sophisticated in their ability to infer insights and predict behaviours based on user data, there is a pressing need for robust privacy regulations, transparent data practices and enhanced user control mechanisms to safeguard individuals’ privacy rights and ensure ethical and responsible AI deployment.¹¹ AI also introduces new privacy concerns for consumers (see Box 3.1), and the use of data as inputs into AI models also raises IP concerns (see Section 3(a)(iv)). As a result, a delicate balance needs to be found between privacy concerns and the need to access large amounts of data to train AI models (see also the opinion piece by Shin-yi Peng).

Restrictions on cross-border data flows also negatively impact trade in AI-enabled products. While there is empirical evidence that AI significantly enhances international trade in digital services, cross-border data regulation can impede such trade. Sun and Trefler (2023) find that restrictions on data flows can reduce the value of AI-enabled apps, making them less attractive to international users. While AI leads to a 10-fold increase in the number of foreign users, the impact of AI on foreign users is halved if the foreign users are in an economy with strong restrictions on cross-border data flows. Thus, economies with strict data regulations may lose out on AI-driven trade opportunities. Striking the right balance between protecting privacy and fostering innovation is therefore crucial for maximizing the benefits of AI for international trade. However, cross-border data flow measures, when aimed at protecting privacy, can help to build trust in AI systems and promote their wider use. A study by OECD and WTO (2024) on the implications of data flow restrictions finds that, although removing data flow regulations across all economies would reduce trade costs, it would also undermine trust, leading to reduced consumer willingness to pay for products and a negative effect on GDP.

(iii) Ensuring the trustworthiness of AI without hindering trade

Standards and technical regulations play a key role in ensuring that AI is trustworthy and, through this, in promoting trade in AI-enabled products. There is growing consensus concerning the pivotal role that regulations, standards and other government interventions can play in ensuring that AI is trustworthy, i.e., that it meets expectations in terms of criteria such as reliability, security, privacy, safety, accountability and quality in a verifiable way.¹³ Ultimately, this means striking a regulatory balance, whereby the benefits of AI are harnessed while its risks are mitigated. Ensuring trustworthiness is not only important for what happens within economies. It is also relevant for what happens outside economies and between borders. Indeed, the internal regulations that governments adopt to protect their consumers can help to build consumers', importers' and other stakeholders' trust in AI-enabled products, thereby fostering trade in such products.

Striking a balance between regulating AI for legitimate policy reasons and enabling trade to flow as smoothly as possible can be particularly challenging. While the challenge of striking the right balance between regulation and free trade is not new, AI's evolving, opaque and multifaceted nature, and the new types of risks associated with it, are making this balancing act in AI regulation and governance particularly complex (see also the opinion piece by Eduardo Paranhos).

Regulating AI requires regulating a product's "behaviour". As mentioned in Chapter 2, "autonomy" is one of the unique attributes of AI. The fact that AI systems can imbue products with various degrees of "autonomy" means that they may generate new forms of risks stemming, not from problems related to the physical components of the product, but instead from the way AI can make the product "behave".¹⁴ Such risks are not easy to foresee, control or even quantify.¹⁵ As Judge et al (2024) note, a unique, defining technical characteristic of AI is that, unlike all other engineered systems, AI's "behaviour" is not dictated or pre-determined by its programme code; it is an "emergent" property. Therefore, AI-enabled products may generate risks for reasons other than those inherent to the tangible elements in the products themselves. For instance, some consider that the "behaviour" of AI-enabled co-bots (i.e., collaborative or companion robots), if unchecked, could provoke mental health problems in the humans they accompany.¹⁶

The opacity of the behavioural nature of AI can make regulation even more challenging. Risky "behaviours" of AI-enabled products may be linked to the way their algorithms are designed. AI algorithms are notoriously opaque (Lim, 2021; Lund et al., 2023). As noted in Chapter 2, transparency and explainability are critical for understanding how and why AI systems work and behave the way they do.¹⁷ This challenge is commonly referred to as the AI "black box" problem,

which has been described as presenting two dimensions, legal and technical. Addressing the legal dimension requires accessing the source code. This may prove difficult as source codes are normally proprietary, i.e., protected by IP, normally in the form of trade secrets. There are, however, regulatory ways to deal with this challenge, for instance, by allowing forced source code disclosure for regulatory or law enforcement purposes,¹⁸ even if in practice this may not be easy or warranted.¹⁹ To some, the technical dimension of the black box problem may be even more significant, as the opacity of an AI system may persist even when access to the source code is free or has been voluntarily or mandatorily granted. Indeed, there may be instances when AI applications are so complex that even programmers themselves are not able to divine an intelligible explanation from the source code and other proprietary information and data as to why and how certain decisions and classifications were reached by the AI system. For some, this means that, until this technical challenge is satisfactorily addressed, regulatory solutions based on open-source disclosure may be "significantly frustrated" (Lin, 2021; Mitchell et al., 2023; Pasquale, 2015).

Adding to the difficulty in pinpointing the source of vulnerability of an AI-enabled product is the fact that their evolving nature may be also triggered by external factors. Such factors include customization: the ability of millions of individuals to "personalize" their AI-enabled products in almost infinite different ways, posing a challenge for regulators to anticipate potential risks associated with each unique customized product. Another factor is connectivity, which may render products vulnerable to cyberattacks or cyberthreats by bad actors that can be located anywhere in the globe. These factors further increase the difficulty for regulators in anticipating and addressing a wide range of possible unforeseeable and unintended risks over the lifecycle of these products (Lund et al., 2023).

AI's dual-use potential may add another layer of complexity. As noted in Chapter 2(a), AI's dual-use nature means that it can be employed for both civil and military purposes. This may add a domestic security and geopolitical dimension to AI's governance, making regulatory interventions and cooperation even more complex (Csernaton, 2024; Klein and Stewart, 2024; Pouget, 2023; Raul and Mushka, 2024). A related issue concerns policy and regulation in the area of AI and cybersecurity (see Box 3.2).

For goods, "traditional" regulations and standards that normally focus on tangible, visible, static product requirements may not be able to address risks stemming from the integration of AI into "traditional" products. The changeability of AI-enabled products, resulting from the evolutionary nature of AI, makes regulation a perennial moving target. AI systems confer new properties and functions to the products into which they are embedded. As stressed in Chapter 2, these products' properties and functions can be described as "dynamic", i.e., they change overtime as a consequence of constant changes occurring throughout the AI system's lifecycle via software updates or other self-improvements resulting from the algorithmic "learning" process. This contrasts with the "static" properties of more traditional products, which normally remain

Box 3.2: AI, cybersecurity and technical barriers to trade (TBT)

AI's ability to analyse large datasets can help in countering cyber threats and responding to malicious cyber-attacks. However, there are concerns related to potentially biased decision-making, the lack of transparency and explainability of AI systems, and potential misuse or abuse. Bad actors can use AI to create new malware, to design new, sophisticated, or targeted phishing attacks, to identify new avenues of attack, and to create deep fakes. Unsurprisingly, cybersecurity is a core concern expressed not only in domestic AI policies but also in international AI principles and governance discussions.

Cybersecurity vulnerability risks are growing as digital technologies are permeating more and more societies and economies. In response, governments are increasingly

adopting cybersecurity-related measures and policies, many in the form of TBT measures, i.e., technical regulations, standards and conformity assessment procedures.

Indeed, cybersecurity-related TBT measures have recently become one of the most prominent digital-technology-related issues discussed in the WTO TBT Committee. To date, more than 90 cybersecurity-related TBT measures have been notified to the Committee, around 65 per cent of these in the last three and a half years. Members have also increasingly raised specific trade concerns (STCs) in the TBT Committee against cybersecurity-related TBT measures: of the 29 STCs raised since 1995, 38 per cent were

raised in the last three and a half years alone.

Cybersecurity was the focus, for the first time, of a specific thematic session of the TBT Committee organized in 2023. Given the global nature of the problem, it was argued in that session that unilateral government interventions in this area should be avoided, as they could ultimately undermine global cybersecurity efforts. The need for governments and the private sector to work in a more coordinated and collaborative manner to address rising regulatory fragmentation and divergence in this area and find better ways to fight increasing cybercrime and cyber incidents was also underscored. In this respect, efforts to develop ambitious, fair and inclusive cybersecurity international standards were highlighted.

essentially the same throughout their lifecycle. Many of the constant changes to properties and functions in AI-enabled products are meant to be beneficial improvements (some even call this “evolution”).²⁰ However, this dynamic process means that known risks and concerns may also be constantly changing, or new ones may be emerging. For AI-enabled products, as is the case for most other products, specifications and requirements will continue to be needed to address risks associated with their “physical” aspects (e.g., hazards from defective mechanical components of an autonomous vehicle). However, for some, such “traditional” specifications and requirements may be ill-suited or insufficient to address situations where the root cause of a risk is not a mechanical or “physical” failure, but an algorithmic design flaw or other problem with the AI system embedded in the product and which may cause it to display risky “behaviour” (e.g., an autonomous vehicle that causes injuries to people or damage to property).

AI-enabled products may cause not only material but also immaterial risks. An AI-enabled product may present both material risks, which are easy to quantify and measure (e.g., physical injuries or damage) and immaterial risks (e.g., privacy or other fundamental rights), which are more difficult to quantify and measure. Material and immaterial risks can sometimes even stem from the same situation.

For instance, in an AI-enabled autonomous vehicle, mechanical malfunctions and/or algorithmic flaws in its internal and external cameras can both cause injuries (material) and affect the privacy of passengers or pedestrians (immaterial).²¹ Likewise, product specifications laid down in one-size-fits-all regulations and standards may be ill-suited for regulating AI-enabled products with different customized solutions (Lund et al., 2023). To address such regulatory challenges, while supporting the deployment of, and trade in, trustworthy AI-enabled products, it has been proposed that regulators think of creative ways to ensure that product requirements and specifications are dynamic and adaptable to the behavioural and evolutionary nature of these technologies, to ensure that they do not become obsolete as AI characteristics, risks and vulnerabilities evolve throughout the product lifecycle.²²

The constantly evolving nature of AI-enabled products may also necessitate new approaches to certify their compliance with regulatory requirements. Indeed, if an AI-enabled product has successfully undergone testing, verification or other certification procedures prior to being placed on the market, this may not necessarily mean that the product will remain certifiable throughout its lifecycle. AI-enabled products, in particular internet-connected IoTs or robotics, may generate new risks after their deployment due to “mutability” factors such as new updates, new data,

unforeseen changes of attributes and functions due to user customization, or unforeseen autonomous behaviours (see Box 3.2).²³ As already discussed, assessing the conformity of some AI-enabled products with underlying technical regulations and standards may also require access to source code, which raises IP-related issues (see Chapter 3(a)(iv)). Regulators may also face challenges in assessing the compliance of AI-enabled products with various novel regulatory requirements that aim, for example, to assess the quality of data used in such products. In light of such a multiplicity of challenges, some consider that regulators may need to re-evaluate their conformity assessment approaches and come up with methods of ensuring effective continuous compliance of ever-changing AI products with underlying technical regulations and standards (Lund et al., 2023; Meltzer, 2023).²⁴

The integration of AI in goods and services has also broadened the scope, number and nature of risks and concerns that regulations and standards need to address. As mentioned above, in addition to “traditional” regulatory concerns, such as interoperability, safety, security, quality, and the protection of human life or health, the use and deployment of AI may also create various “non-typical” risks, that some even qualify as “existential” (UNDRR, 2023), and may raise complex ethical and societal questions affecting public morals and human dignity.²⁵ If AI is trained on biased and skewed datasets, it may perpetuate or exacerbate biases or discrimination against minority groups and infringe upon individual rights and freedoms (see Chapter 2). AI-enabled goods and services are also a cause for significant concern with regard to data privacy, as they involve the collection, processing and storage of vast amounts of user data (see Chapter 3(a)). In addition, as mentioned above, AI is a technology prone to dual use, which may raise complex geopolitical and domestic security issues and lead to further regulatory fragmentation. Finally, both AI “inputs” and “outputs” raise new and complex issues of IP protection and ownership (see Chapter 3(a)(iv)).

These concerns render it challenging to design proper regulatory solutions to ensure the trustworthiness of and support trade in AI and AI-enabled products. As already mentioned, AI raises societal and ethical concerns (“immaterial” risks) that, unlike “traditional” concerns such as health and safety (“material” risks), are not typically a subject for technical regulations and standards. Such “non-technical” concerns are more difficult to regulate, monitor and enforce compared to more traditional regulatory objectives, such as product safety or the protection of human health or life, which can be addressed in more “technical” and objective ways. It has been argued that AI governance and regulatory frameworks may require norms, regulations and standards that are perhaps better described not as purely “technical”, but instead as “socio-technical” instruments, i.e., combining technical issues with broader societal considerations (Dentons et al., 2023; Kerry, 2024; Meltzer, 2023). Pouget (2023) argues that developing socio-technical regulations is challenging in situations where both the technology and the harms it can cause are “so complex that it becomes

difficult to separate value judgements from technical detail”. Some have even questioned whether this could ever be done in practice.²⁶

Such non-typical or immaterial AI-triggered risks and concerns may also be intrinsically prone to regulatory fragmentation, which could hinder trade. Indeed, it might be difficult for legislators to agree on common international denominators with respect to some AI-related societal values and concerns such as ethics, privacy or human rights, the relative importance of which may vary across economies.²⁷ Unnecessary or avoidable regulatory fragmentation could, in turn, hamper the opportunities and benefits associated with AI (Bello Villarino, 2023; OECD, 2022a). In particular, it could result in high regulatory compliance burdens and costs, and consequently create non-tariff barriers to trade for AI businesses.

(iv) How AI is shaped by and may reshape IP

AI poses new conceptual challenges for the traditional, human-centric approach to IP rights. Balanced IP rights and their enforcement have an important role to play in ensuring both equitable access to AI technology and a fair distribution of economic gains from its use. AI raises several important questions in this respect.

A first question concerns what form of IP protection AI algorithms are granted. If the IP protection is based on the fact that these algorithms are trade secrets – and thus that secrecy is an essential requirement on which to establish IP protection – this raises issues concerning a lack of transparency. Alternatively, new and inventive algorithms may be protected by patents in some jurisdictions, with the patent system’s mandatory disclosure mechanism yielding extensive information about AI technologies, which directly passes into the public domain in many economies (WIPO, 2024). However, patent protection may constrain development of algorithms in economies in which patents have been taken out. Copyright, another type of IP protection, can be automatically extended to both source and object code, which may constrain analysis and use of algorithms. As envisaged by the objectives of the international IP system, appropriate exceptions and limitations to IP rights protection are needed to balance the different interests and to ensure appropriate access and dissemination of AI technology. These regulatory tools may have to be adapted for this specific context, and some jurisdictions have taken legislative steps or developed policies to encourage the development of open-source AI technologies.²⁸

A second question concerns the use of copyright-protected data as AI inputs. Under the current international IP legal framework, materials such as original texts, images and compilations of data may be subject to copyright protection. This may raise the question of whether

Opinion piece

Navigating AI regulation: balancing innovation, risks and regulatory defragmentation

It is not a simple task to determine when a new social or economic phenomenon warrants regulation. Those challenges can be further amplified when the new scenarios take the form of innovative technologies, posing both risks and opportunities. In addressing this, regardless of the nature of the changes, it is important to reflect on a few foundational questions: (i) what risks and opportunities are at stake; (ii) how well understood those new technologies are, so that the tools to tackle the possible risks can be properly balanced; and (iii) which aspects of the technological progress indeed require new rules, vis-à-vis the existing laws.

AI is transforming the way we work, communicate and create content faster than ever before. Fostering the development and implementation of AI solutions has the potential to increase efficiency and job quality, as noted, for example, in recent studies by the International Labour Organization and the consultancy firm McKinsey. Yet, for these benefits to materialize, we should consider which traits of the AI systems should really be regulated and aim to establish a model that, at the same time, is capable of protecting society and promoting – rather than discouraging – research and development.

It is an oversimplification to say that economies are mostly weighing up either context-principle based formats to regulate AI, or prescriptive models with a more detailed set of obligations and sanctions.

In Brazil, the debates around AI regulation have so far examined aspects from each of those possible structures: a prescriptive model, openly inspired by the European regime, and another proposal for a context-based framework anchored on widely recognized principles for governance and risk mitigation – e.g., those of the United Nations Educational, Scientific and Cultural Organization (UNESCO), the Organization for Economic Co-operation and Development (OECD) – reaffirming the role of existing legislation.

Indeed, it is possible that most situations raising concerns about the deployment of AI could be dealt with through existing federal laws, particularly in

Eduardo Paranhos

Lawyer, LL.M – University of
London (LSE); Chevening Scholar;
Head of AI Working Group at Brazilian
Software Association (ABES)

relation to privacy, product safety, consumer protection and internet regulations, as well as some downstream regulations, such as Brazil's health authority ordinance on software as a medical device (SaMD).²⁹ For instance, a data breach that occurs as a result of using an AI system is subject to the same controls and remedies provided for in Brazil's privacy law as other breaches that occur without the use of AI. The privacy law also covers the potential misuse of sensitive personal data that generates biased outputs in the same way as similar misuse in offline settings.

However, it is of pivotal importance to map the gaps in the current legislation clearly, so that fresh AI regulations can address that very gap, avoiding overlaps and the resulting legal uncertainty. Another critical point which could make prescriptive models problematic is the emphasis on regulating the “development” of AI, instead of focusing on high-risk “uses” of the technology.

One final consideration refers to the level of preparedness that the upcoming regulations should display to be able to evolve with the technology. It seems more realistic that context-based regulations should gradually progress into stricter forms “if” and “when” needed, than the other way round. Similarly, in the context of trade-related concerns over the adverse effects of regulatory fragmentation, AI regulations that concentrate on high-risk uses – not on the development of the technology – could facilitate the pursuit of a more harmonized approach across markets. Notably, some economies that are global protagonists in investments and implementation of AI³⁰ have been leaning towards evolutive regulatory formats, balancing AI governance and risk mitigation goals, while helping to raise living standards, creating quality jobs and improving people's lives through responsible innovation.

Disclaimer

Opinion pieces are the sole responsibility of their authors. They do not necessarily reflect the opinions or views of WTO members or the WTO Secretariat.

their use in training AI amounts to copyright infringement. This question translates into whether such use is now, or should be, automatically permissible under exceptions to exclusive copyright (e.g., for educational use of copyright-

protected materials). The application of such limited exceptions to generative AI, in particular, is more complex than in traditional cases, due to factors like the scale of data used and the purpose of the use.

A third set of questions concerns whether AI outputs generated autonomously can be subject to IP protection. An AI output is created based on the patterns and rules learned during the training process, by means of the input data. However, this output (including in the form of content) is not a mere reproduction or recombination of this material. Rather, the output can take the form of novel and creative material, reflecting the AI system's ability to understand and mimic the complexities of human-generated content. Thus, AI output may encompass a wide range of creations and innovations,³¹ including artwork, literary works, music, design, films, video games and inventions. As AI is increasingly capable of producing outputs autonomously, the lines between human and AI contributions to creation or inventions are increasingly becoming blurred, making the question of inventorship and authorship more pressing and complex.

Various approaches have been taken, or proposed, for finding balanced and equitable answers to some of the above questions, both in terms of AI inputs and outputs. In terms of AI inputs, proponents of the use of copyrighted material to train AI argue that this constitutes a “transformative” use, as the model does not replicate the copyrighted works, but instead generates new content inspired by the learned patterns. They also argue that such use does not negatively affect the market for the original works and that, in some cases, it could potentially complement that market. This is still an ongoing debate. A balanced approach taking into account both the moral and the economic interests of creators of original works and users of AI needs to be found, and the legal community continues to explore these issues. Approaches to this issue differ significantly across jurisdictions (see Chapter 3(b)).

The question of the protection and ownership of AI generated outputs necessitates a re-evaluation of existing IP legal frameworks. As noted above, AI may generate outputs in the form of a work or invention. This in turn raises questions about whether and/or under which circumstances IP rights can be granted for AI-generated creations or innovations, and if so, who, if anyone, owns the resulting IP, or who is liable if the output violates the IP of others. Is it the developer of the AI tool, the user who prompted the AI output, or neither, since the creator or inventor is not human? Current IP laws attribute authorship and inventorship, as well as resulting economic rights, to humans. In some jurisdictions, AI itself is not recognized as the creator or inventor within the current IP legal framework.³² In other jurisdictions, the issue is open to judicial interpretation based on existing laws.³³ With the widespread deployment of AI, which can now be used by individuals across the globe, the question of protection and ownership of AI generated outputs becomes a constant and global concern. This question not only challenges the traditional understanding of creativity and ownership, but also urges a re-evaluation of existing IP legal frameworks in the age of AI (see also Chapter 4(f)).

(b) The global race to promote and regulate AI and the risk of fragmentation

The immense potential of AI has prompted governments around the globe to take action to promote its development and use while mitigating its potential risks. However, the increasing number of domestic, regional and international initiatives and their design are fragmenting the policy landscape, with possibly negative consequences for companies trading internationally. The economic costs of fragmentation highlight the importance of mitigating regulatory heterogeneity.

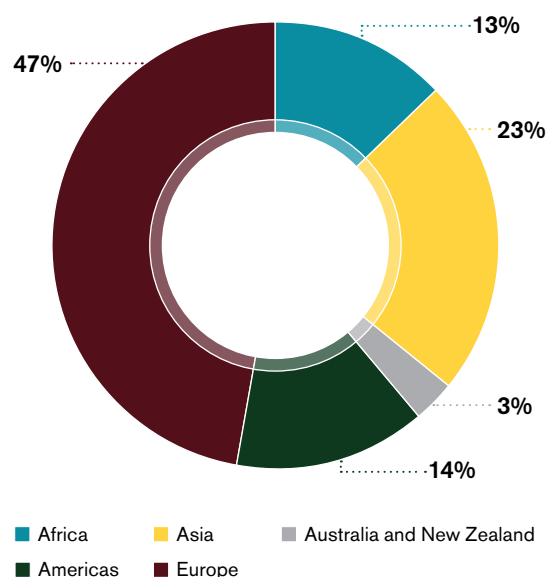
(i) Domestic initiatives

Governments are using a variety of instruments to promote AI and to address and mitigate its risks. These range from AI-specific strategies and policies to sector-specific legislation (including data regulations) and trade policy measures. However, there are already signs that heterogeneity in the design of these measures may be leading to regulatory fragmentation. The sheer number of domestic strategies and policy initiatives related to AI indicates that AI is an area of priority and that a sustained high level of intervention can be expected in the near future, with a potential risk of growing regulatory fragmentation.

AI strategies and policies

An increasing number of jurisdictions is putting in place AI strategies and policies at the domestic level. The number of economies that has implemented AI strategies³⁴ increased from three in 2017 (Canada, China and Finland) to 75 in 2023. Canada initiated the first domestic AI strategy in March 2017. In 2023 alone, eight new strategies were added by economies in the Middle East, Africa and the Caribbean, showcasing the worldwide expansion of AI policymaking (Maslej et al., 2024). In addition, or as part of domestic AI strategies, governments around the world have taken over 1,000 AI policy initiatives.³⁵ The majority of AI policy initiatives are concentrated in Europe, followed by Asia, the Americas and Africa (see Figure 3.4). Most AI-related legislation passed since 2016 aims to enhance an economy's AI capabilities, such as establishing a network of publicly accessible supercomputers, as opposed to restrictive legislation, which imposes conditions or limitations on AI deployment or usage (Maslej et al., 2024).

Figure 3.4: Share of AI policy initiatives by region (2024)



Source: OECD database of domestic AI policies (<https://oecd.ai>).

The European Union, a WTO member in its own right,³⁶ has been particularly active, with the adoption of a series of policy measures to support the development of trustworthy AI at the EU level.

Policy measures to support the development of trustworthy AI include the AI Innovation Package,³⁷ the Coordinated Plan on AI,³⁸ the “Proposal for standard contractual clauses for the procurement of Artificial Intelligence (AI) by public organisations”, and the EU AI Act (AIA) (European Union, 2024). The AIA, which was formally adopted in 2024, is the world’s first comprehensive horizontal legal framework on AI.³⁹ The main stated objective of the AIA is to ensure that AI systems within the EU are safe and comply with existing laws on fundamental rights, norms and values.⁴⁰ The AIA adopts a risk-based approach to regulating AI systems.⁴¹

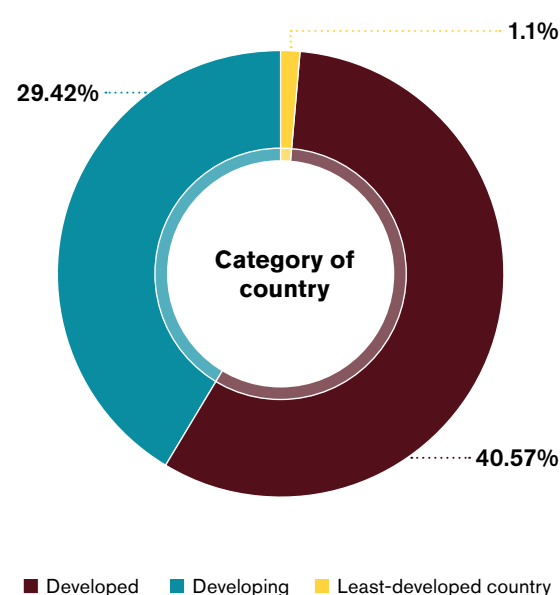
Most domestic AI policy initiatives are implemented by developed economies, reflecting the growing AI divide.

While a reasonable share (around 30 per cent) of developing economies have put AI policy measures in place, only one LDC, Uganda, has done so, with two sector-specific policies and one general policy on AI governance (see Figure 3.5). The increasing attention being paid to AI in policymaking can also be seen in references to AI in legislative proceedings, which have increased almost tenfold across the globe since 2016, and nearly doubled between 2022 and 2023 (Maslej et al., 2024).

Domestic AI policy initiatives can be classified into four broad categories: governance, financial support, guidelines and regulations, and AI enablers. As of

March 2024, more than a third (36 per cent) of AI policy initiatives listed by the OECD Artificial Intelligence Policy Observatory⁴² concerned governance of AI. Governance aspects usually focus on the establishment of frameworks for AI development and deployment, including vertical and horizontal coordination, AI’s integration into public sector, public consultation and evaluation mechanisms, and the creation of regulatory bodies or committees to oversee AI-related activities. Close to 19 per cent of domestic AI policy initiatives aim to provide financial support and incentives for AI research, development and adoption.⁴³ Around 18 per cent of these initiatives include the development of guidelines and regulations (on issues such as data privacy, algorithmic transparency, bias mitigation and safety standards to promote the responsible and ethical development and use of AI technologies), the establishment of regulatory oversight and ethical advice bodies to provide guidance and supervision in navigating these regulations effectively, and the development of standards and certification processes to facilitate the development and adoption of AI technologies in compliance with regulatory requirements and ethical principles. Finally, a large number of domestic AI policy initiatives (around 27 per cent) focuses on fostering an environment conducive to AI innovation and adoption (i.e., AI enablers). These include initiatives to enhance AI-related skills and education to attract talent, public awareness campaigns to raise awareness about AI, and the establishment of collaborative platforms, to bring together stakeholders within the innovation ecosystem, and business advisory services, to support innovation and entrepreneurship (Maslej et al., 2024). Some of these initiatives, like the 2023 US “Executive Order on the

Figure 3.5: Share of AI policy initiatives by level of development (2024)



Source: OECD database of domestic AI policies (<https://oecd.ai>).

Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence”, contain measures to support AI-related hardware, such as computing infrastructure, as well as competition and innovation in the semiconductor industry.

Governments seem to be preparing or adopting an increasing number of detailed rules and regulations related to implementing and enforcing AI legislation.

According to Stanford University's 2024 “AI Index”, the number of AI-related regulatory measures has risen significantly in the United States and the European Union over the past few years. There were 25 AI-related regulatory measures adopted in the United States in 2023, including three related specifically to international trade and international finance, compared to just one in 2016. The total number of AI-related regulatory measures grew by 56.3 per cent in 2023 alone to reach 83. As for the European Union, it has passed almost 130 AI-related regulatory measures since 2017, including 13 led by the Directorate-General for Trade and the Directorate-General for Competition (Maslej et al., 2024). Several economies are also developing strategies or putting in place specific initiatives to develop AI standards (see Box 3.3).

Environmental concerns are currently high on the policy agenda. Governments are therefore also starting to draft regulatory frameworks to address the potential negative environmental impacts of AI and to harness its many benefits. For example, one of the EU's

AIA objectives is to assure the “environmental protection against harmful effects of [AI] systems in the Union and supporting innovation.”⁵⁴ See also Box 2.1 on AI's environmental impacts.

An increasing number of jurisdictions are also putting in place AI-related “sandboxes”.

The objective of these is to test new economic, institutional and technological approaches and legal provisions under the supervision of a regulator for a limited period of time.⁵⁵ About a dozen jurisdictions, including Colombia, Estonia, the European Union, France, Germany, Lithuania, Malta, Norway, Singapore and the United Kingdom have such structures in place (OECD, 2023).

Some jurisdictions are also developing “GovTech” tools (digital tools used to optimize public services) to address the new regulatory challenges raised by AI and to promote trustworthy AI.

A notable example is Singapore's “AI Verify” tool, developed by the Infocomm Media Development Authority and Personal Data Protection Commission.⁵⁶ AI Verify is an open-source software tool to assess the trustworthiness of AI systems according to a set of criteria and factors. The tool, which is at minimum-viable-product stage,⁵⁷ aims to automate transparency assessment of AI systems, which would allow companies to see whether new AI systems comply with relevant international standards and regulations (see the case study on Singapore's approach to AI in Box 3.4).

Box 3.3: Domestic standards on AI⁴⁴

Standards play an important role in domestic AI policy approaches and several economies are developing strategies or putting in place specific initiatives to develop AI standards.⁴⁵ Some economies even recognize AI as one of the priority areas in their general standardization strategies.⁴⁶

As of July 2024, almost 170 standards are being developed or have already been published by various domestic standards-setting bodies (such as BSI, CEN, CENELAC, NIST).⁴⁷ Most of such domestic standards seem to be of horizontal application, while others seem to be sectorial, i.e. only covering specific industries and sectors such as transportation, healthcare or energy. AI-related standards cover

a variety of broad topics, most often those related to data management, quality, processing and protection, as well as risk management, safety and security, interoperability, and organizational governance. These standards address specific technical requirements such as process, management and governance, measurement and test methods, terminology, interface and architecture specifications, and product and performance requirements.⁴⁸

One common feature across various domestic standardization approaches is the recognition of the importance of engagement and cooperation on AI standardization at the international level (Kerry, 2024).⁴⁹ For instance, Australia's AI Action Plan reflects

Australia's intention to participate in international standards-setting processes,⁵⁰ while China's Global AI Governance Initiative encourages international cooperation for developing AI standards based on broad consensus.⁵¹ In the same vein, the US Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence mandates relevant agencies to cooperate with standards development organizations to drive the development of AI-related consensus standards.⁵²

In this respect, as in other regulatory areas, domestic standardization efforts on AI will tend over time to rely on international standards-setting work.⁵³

Box 3.4: Case study: Singapore's approach to AI

For Singapore, AI is a necessary means to overcome natural constraints, such as a small labour force on a small landmass, and to raise the productivity and strengthen the competitiveness of its industries, both in globally tradable sectors, such as trade, finance, and in domestic services, such as retail and food and beverages.

For example, Singapore has leveraged AI in order to continue to act as a global hub facilitating trade and connectivity. Singapore's Changi Airport, which handled more than 59 million travellers last year, uses AI to screen and sort baggage, and to power facial recognition technology for seamless immigration clearance. The Port of Singapore, which handled cargo capacity of 39 million twenty-foot equivalent unit (TEUs) in 2023, uses AI to direct vessel traffic, map anchorage patterns, coordinate just-in-time cargo delivery, process registry documents, and more. To facilitate communications and business exchanges across a linguistically diverse region of 680 million people who speak over 1,200 different languages, Singapore has also invested in developing the world's first large language model tailored to Southeast Asia's languages and cultures; this open-source model is dubbed SEA-LION, short for Southeast Asian Languages in One Network.

In 2019, Singapore issued a framework for responsible AI use. The Model AI Governance Framework provides detailed and practical guidance to address key ethical and governance issues when deploying AI solutions. In 2024, Singapore further extended the Model Framework beyond traditional AI to address generative AI and the novel risks it poses. Within the Association of Southeast Asian Nations (ASEAN), Singapore has spearheaded the development of an ASEAN Guide of AI Governance and Ethics. At the United Nations, Singapore convenes the Forum of Small States (FOSS), a grouping of 108 small economies, and introduced a Digital Pillar in 2022, which provides baseline capacity-building for issues including AI, most recently through the AI Playbook for Small States, which was co-developed with Rwanda and launched at the UN Summit of the Future in September 2024.

Singapore works closely with a range of partners, bilaterally and in various groupings, on guidelines for AI developments and innovations. With the United States, Singapore has deepened information-sharing and consultations on international AI security, safety, trust and standards development through collaborations in AI, including the US-Singapore Critical and Emerging Technologies Dialogue.

With China, Singapore is enhancing mutual understanding of approaches to AI governance, such as under the inaugural Singapore-China Digital Policy Dialogue. Singapore also participates in the G7 Hiroshima Process, the AI Safety Summit series, the OECD AI Principles, the Global Partnership on AI (GPAI) and the World Economic Forum's AI Governance Alliance.

In 2022, Singapore launched AI Verify, an AI governance testing framework and a software toolkit, which contains baseline standardized tests, covering core principles of fairness, explainability and robustness. In 2024, Singapore launched AI Verify Project Moonshot, which broadens the original toolkit to cover generative AI and return intuitive results on the quality and safety of large language models. Given that the science of AI testing and governance is still nascent, Singapore has also set up the AI Verify Foundation to harness the collective power and contributions of the global open-source community to jointly develop AI Verify testing tools. The Foundation has grown to more than 110 members and includes companies such as Google, IBM, Microsoft, Red Hat, Meta and Salesforce.

Source: Based on inputs from the Ministry of Digital Development and Information, Singapore.

The heterogeneity of domestic initiatives may lead to unintended fragmentation. Analysing eleven AI rulebooks from seven jurisdictions (i.e., Argentina, Brazil, Canada, China, the European Union, the Republic of Korea and the United States), Fritz et al. (2024) find that governments prioritize different objectives with their AI regulation, use substantially different regulatory requirements to achieve the same priorities, and choose different scopes and formulations to achieve the same regulatory requirement for a shared priority, leading to unintended fragmentation at the level of priority, requirement and scope.

Unintended fragmentation extends to non-AI-specific, sector-specific legislation, such as AI-relevant IP and data regulations. Approaches to copyright "fair use", for example, differ significantly across jurisdictions (see Chapter 3(a)(iv)). While Japan modified its Copyright Act in 2018 to allow machine learning models to use copyrighted works for any purpose, including commercial use, without needing explicit permission from copyright holders,⁵⁸ the EU AIA is much less permissive. According to the EU AIA, the provider of a generative AI model, whether open-source or closed, must establish a policy to respect EU copyright law,

Box 3.5: The challenge of navigating AI regulations: The case of Canvass AI

Invited to speak at a WTO workshop on regulatory cooperation on digital products, Humera Malik, CEO of Canvass AI, a startup that provides industrial AI solutions to enhance operational efficiency, profitability, and sustainability, explained that the diverse approaches

to regulate AI make achieving global reach difficult. Divergent regulations strain resources, make the navigation of rules without specialized knowledge difficult, and impact market entry. Moreover, data protection regulations impose additional restrictions, affecting AI development and cross-

border market entry. She added that “minimizing complexity and promoting convergence would greatly ease compliance efforts”.

Source: https://www.wto.org/english/tratop_e/tbt_e/tbt_2006202310_e/tbt_2006202310_e.htm.

including the EU Directive 2019/790 on Copyright and Related Rights in the Digital Single Market (CDSM). Under the CDSM, research organizations are permitted to reproduce and extract copyrighted works for text- and data-mining purposes without requiring the authorization of the copyright-owner, provided that these research organizations have lawful access to the works, and that the use is for the purposes of scientific research. The use of copyrighted materials for text- and data-mining for any reason is also permitted beyond scientific research, but in this context, copyright-owners have the option explicitly to reserve their rights and thereby prevent the use of their works for text- and data-mining without their approval (European Parliament, 2024). Similarly, the AI Bill pending adoption in the Brazilian Congress provides, for example, for a limited copyright exception when the extraction, reproduction, storage and transformation taking place in data- and text-mining processes are carried out by research and journalism organizations and institutions, museums, archives and libraries. As for the United States, while there is still no legislation or regulation on this issue, a high profile ongoing litigation case was filed by the New York Times against OpenAI for the unauthorized use of its content in December 2023. Another example is the diverging approaches to algorithmically authored works (see Chapter 3(a)(iv)). While the United Kingdom protects algorithmic creations, albeit without recognizing AI itself as an author,⁵⁹ Australia and the United States make it clear that a human author is needed (Liu and Lin, 2020). Finally, some jurisdictions provide expansive protection to trade secrets, applying proprietary protection to source code, algorithms, training materials and datasets used to train AI models, while others do not provide them with exclusive IP protection (Kilic, 2024). Beyond IP, data regulations are also marked by a high level of fragmentation.

The design of some measures may affect market competitors in other economies and have trade-distortive effects, leading to further fragmentation. The significant economic potential of AI is leading political

authorities in various jurisdictions to put in place measures to promote the development of AI. These include the creation of “AI factories”, to give AI start-ups and small businesses access to supercomputers on which to build their own models, research initiatives to connect researchers and educators to computational, data and training resources to advance AI research and research that employs AI, and subsidies for firms that purchase domestically produced AI chips. Some of these measures appear to limit opportunities to domestic entities or to provide incentives on the condition that domestic products are used (Aaronson, 2024b).

The economic costs of regulatory fragmentation highlight the importance of mitigating regulatory heterogeneity. The impact of fragmentation can be felt at various levels, including lost trade opportunities, diminished productivity gains and stifled innovation, with potentially important economic consequences for vendors of AI-enabled goods and services (Fritz and Giardini, 2024). As AI technologies become increasingly embedded in goods and services across a wide range of sectors, in the absence of efforts to mitigate regulatory heterogeneity, the resulting costs and other negative impacts are likely to grow significantly. The impact is likely to be particularly important for small businesses, which are already struggling to navigate through divergent regulatory approaches on AI (see Box 3.5).

Data regulations

Regulating data stands high on policy agendas. With the rise of digital technologies, including AI, initiatives promoting access to data to foster domestic innovation and competition, protecting privacy and controlling the flow of data across borders stand high on policy agendas. However, what is emerging is a landscape of measures that is not only fragmented, but that may also have trade-distortive impacts beyond fragmentation.

Open government data and data-sharing initiatives to foster innovation and competition

An increasing number of jurisdictions is taking initiatives to promote open government data to foster business creation and innovation and to increase competition in domestic markets. Recognizing the value of data as a public good, some jurisdictions, both in developed and developing economies,⁶⁰ are pursuing open government data initiatives to promote business creation and innovation and stimulate the domestic digital and AI economy by encouraging the use, reuse and free distribution of government datasets under open data licences. Examples include the EU's Open Data Directive, India's Open Government Data platform⁶¹ and Singapore's "Smart Nation" initiative.⁶² These initiatives come in addition to *ex ante* competition regulations put in place in some markets to better address competition issues raised by the digital economy.⁶³ Open government data is a goal that is also being pursued at the regional and international levels, including in the context of the WTO Joint Statement Initiative on E-commerce (see Chapter 4(a)(v)).

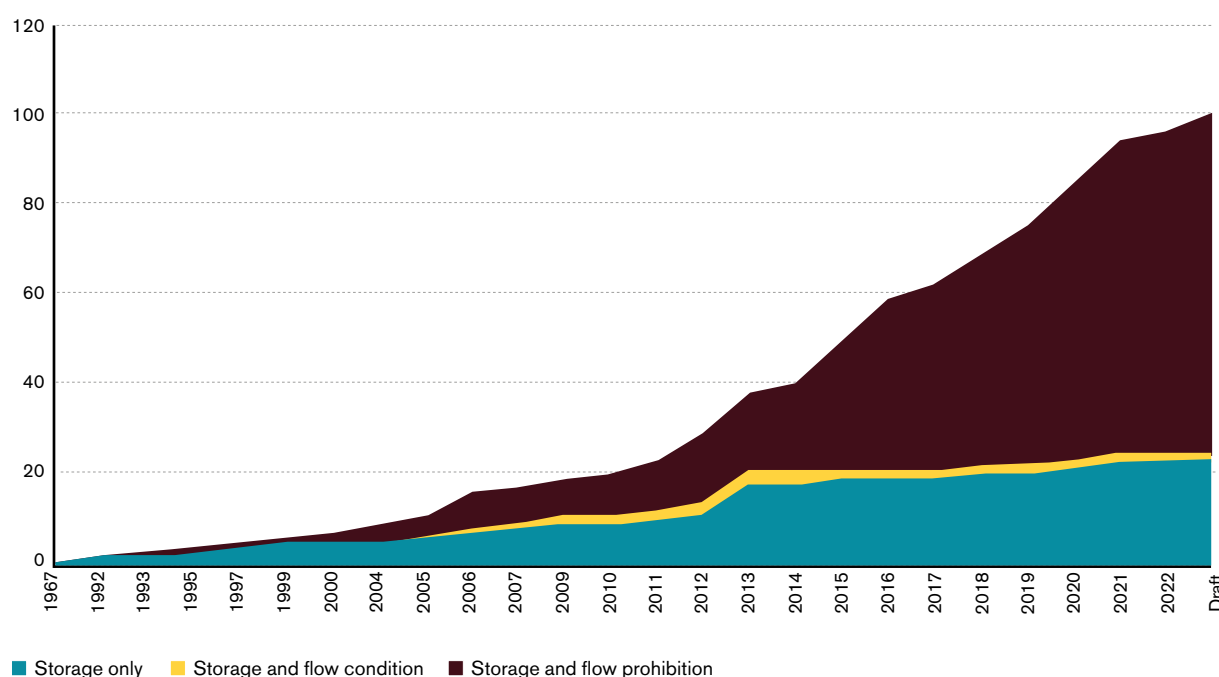
Other approaches aim to promote data-sharing across sectors to foster innovation or to mandate it to counterbalance winner-takes-all dynamics in the digital economy. The EU Data Governance Act (DGA), for example, seeks to increase trust in data-sharing and data availability. It entered into force in 2023 and supports the setup of trustworthy data sharing systems – called Common

European Data Spaces – in strategic domains, involving both private and public players. Training AI systems is listed as one of the key benefits of the initiative (European Commission, 2024b). The EU Data Act, which entered into force in January 2024, complements the DGA and creates the processes and structures to facilitate data-sharing by companies, individuals and the public sector (European Commission, 2024a). The Act protects EU businesses in data-sharing contracts from unfair contractual terms that may be imposed unilaterally by one contracting party on another; the aim is to enable small businesses, in particular, to participate more actively in the data market. Other economies that have put in place data-sharing initiatives include Colombia, Japan and the Republic of Korea. Some jurisdictions, such as Australia and the European Union, are also experimenting with legally mandated data-sharing to foster a competitive environment in which AI startups also have access to large datasets (Mayer-Schönberger and Ramge, 2018; Prüfer, 2020).

The extent to which open government data and data-sharing initiatives support innovation and level the playing field both within and across economies remains unclear. There are concerns that such initiatives may in fact disproportionately benefit large AI firms, as these have the capacity to collect open data and to correlate it with the "closed data" they possess and control to generate new data. As a result, large AI firms stand to gain more than those who lack such capabilities and have to rely on open data entirely, which could amplify the growing AI divide

Figure 3.6: Data localization is growing and becoming more restrictive

(number of measures)



Source: Del Giovane et al. (2023).

between companies (see Chapter 3(a)). Such policies could also have geopolitical implications, as those operating out of relatively big, closed digital economies are able to capture open data elsewhere in addition to the data they collect domestically without much external competition, which could result in further imbalances across economies (Streinz, 2021).

In addition, while some-data sharing initiatives are clearly open to foreigners, uncertainty remains concerning other initiatives. These could raise potential most-favoured-nation (MFN) issues and result in trade-distortive effects. Japan, for example, announced in 2024 that its data spaces would be open to foreigners, but the programmes of some other jurisdictions seem designed to support data-sharing within the jurisdiction concerned, which could have a trade-distortive effect (Aaronson, 2024).

Privacy and data protection

Over the last decades, many governments have enacted regulations for personal data protection to address growing concerns over privacy. According to UN Trade and Development (UNCTAD), more than 70 per cent of jurisdictions – 137 out of 194 – adopted legislation to secure the protection of data and privacy in 2021, with significant differences across levels of development (UNCTAD, 2021a). The share of jurisdictions having passed such legislation is lowest in LDCs (48 per cent). The most well known of these is the EU's General Data Protection Regulation, which became effective in May 2018.

AI raises new privacy concerns for individuals and consumers. This is leading to an increasingly complex trade-off between the need to access large amounts of data to train AI models and privacy concerns. As seen in Chapter 2, AI's reliance on large amounts of data, including personal data, and its capacity to process and analyse vast datasets and to correlate data can lead to privacy breaches and information spillovers, introducing new privacy challenges.

Privacy and personal data protection regulations differ markedly across jurisdictions, affecting the flow of data. Most governments have introduced data protection laws, but these regulations vary significantly from one jurisdiction to another. Whereas some economies, like the United States, primarily rely on the industry to self-regulate the protection of personal data, others follow different approaches that focus on state intervention to defend state sovereignty, citizens' rights, security or domestic development (Bradford, 2023; Jones, 2023; Mitchell and Mishra, 2018; UNCTAD, 2021b). These include limitations on the international transfer of personal data, aimed at maintaining jurisdictional oversight. These different approaches to data governance are creating distinct "data realms" that are fostering a new digital divide between these jurisdictions and others that are rule-takers, creating regulatory uncertainty and barriers to the flow of data across borders (Aaronson and Leblond, 2018; Jones, 2023). The divergence in regulatory approaches between the European Union and the United States has been a particular case in point,

with two data privacy agreements brought down by the European Court of Justice of the European Union.⁶⁴

Cross-border data flow restrictions and data localization requirements

Cross-border data flow restrictions aim to limit the flow of data, and measures to control where data is stored or processed are on the rise. Motivations behind cross-border data flow restrictions and data localization requirements (i.e., explicit requirements that data be stored or processed domestically) vary, ranging from concerns over sensitive data, related to national security, to privacy considerations. Such measures are sometimes seen as an incentive to boost local competitiveness (Aaronson, 2024b; McKinsey, 2022). By early 2023, there were 96 data localization measures across 40 economies in place, with nearly half of the identified measures having emerged after 2015 (see Figure 3.6). Not only has the number of data localization measures increased, but the measures themselves are also becoming more restrictive, with more than two-thirds of identified measures involving not only a storage requirement but also a prohibition for data to flow from one economy to another (Del Giovane et al., 2023). These data regulations apply to different types of data, including personal data, and to different sectors. As noted in Chapter 3(a), striking the right balance between fostering AI innovation through access to data and protecting privacy is crucial for maximizing the benefits of AI for international trade.

The global fragmentation of data flow regulations underscores the need for increased international cooperation. While there are legitimate reasons for diversity in regulation, the current landscape is increasingly complex and fragmented, imposing additional costs on firms, especially those located in small markets, creating uncertainty, and hindering the cross-border flow of data that plays such an essential role in AI development and innovation, in particular for small economies. The economic costs of the fragmentation of data flow regimes along geo-economic blocks are potentially sizeable, amounting to a loss of more than 1 per cent of real GDP, according to an OECD-WTO study (OECD and WTO, 2024). A global approach that balances the need for robust data oversight and protection of privacy, while ensuring that data can be accessed and can flow freely across borders, is needed (Jones, 2023).

Border measures

Many of the hardware components and raw materials crucial to AI systems face increasing export restrictions. Export restrictions applied to industrial raw materials, many of which play a critical role in the manufacturing of advanced chips needed to power AI systems and in communications equipment, increased more than five-fold between 2009 and 2020 (OECD, 2023b). More recently, the race to dominate AI development, combined with broader economic, geopolitical and security considerations linked to the dual-use nature of AI systems, has led a growing number of advanced economies to impose export restrictions on

advanced chips central to AI systems and on the tools used to manufacture them.⁶⁵ In reaction, China, one of the main targets of these measures, requested consultations under the WTO Dispute Settlement Understanding (DSU) in December 2022⁶⁶ and imposed export restrictions on two metals used in chipmaking and communications equipment in July 2023.

There is a risk that these restrictions will affect the global development and deployment of AI technologies and increase economic and, potentially, technical fragmentation. In the short term or when limited alternatives are readily available, restrictions can impact access to the technology by importing economies. A longer-term effect may be that new technological developments will be postponed due to a lack of access to advanced technology, compounding risks of economic and technical fragmentation.

(ii) Bilateral and regional cooperation initiatives to address AI

The increasing number of bilateral and regional cooperation initiatives on AI governance focusing on different priorities adds to the risk of creating multiple fragmented approaches.

Bilateral cooperation initiatives that touch upon issues relevant to AI and trade prioritize different issues. Cooperation between the United States and the European Union in the context of the Trade and Technology Council (TTC), which was established to promote EU-US cooperation, focuses primarily on aligning terminology and taxonomy and on monitoring and measuring AI risks (NIST, 2021). The first AI-related outcome of the TTC was the launch, in December 2022, of a Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management. The Joint Roadmap aims to guide the development of tools, methodologies and approaches to AI risk management and trustworthy AI, to develop a common understanding of key terms, to support and lead development of international standards, and to monitor and measure existing and emerging AI risks. In May 2023, the TTC adopted the “EU-US Terminology and Taxonomy for Artificial Intelligence – First Edition”, which builds on existing standards such as International Organization for Standardization (ISO) standards to define key AI-related terms, but does not define AI.⁶⁷ In April 2024, the European Union and United States launched a Research Alliance in AI for the Public Good. The Research Alliance aims to foster scientific cooperation to better harness AI for the benefit of the environment, energy optimization, disaster reduction and emergency responses.⁶⁸

Other bilateral initiatives in which the United States is involved focus more on collaboration to promote alignment in general terms. In October 2023, the United States and Singapore launched a Critical and Emerging Technology Dialogue,⁶⁹ which establishes a bilateral

AI governance working group focused on advancing shared principles for safe, trustworthy and responsible AI innovation, and calls for strengthened collaboration through joint research and educational funding and for exploring reciprocal certification programmes for American and Singaporean AI professionals on the basis of shared standards, tests and benchmarks. The Dialogue includes cooperation on standard development and a mapping exercise between domestic standard-setting bodies to align approaches. And in April 2024, the United States and Uruguay signed a Memorandum of Understanding (MoU) to foster cooperation on certain critical and emerging technologies, such as semiconductors, AI, data flows, telecommunications and cybersecurity, including by identifying opportunities to support the development and use of relevant international standards and by encouraging interoperability and global compatibility, as well as greater cooperation in multilateral and international organizations.

China’s bilateral initiatives prioritize AI safety and governance, as well as development issues. Dialogue between China and the United States primarily focuses on AI safety and governance. Announced in November 2023, the first dialogue took place in May 2024 (The White House, 2023b; Murgia, 2024). Global governance of AI is also of high importance in China’s discussions with African leaders in the context of the China-Africa Internet Development and Cooperation Forum. The last forum, which took place in April 2024, called for more representation of developing economies in the regulation of AI. Beyond the above-mentioned examples, agreements to maintain bilateral dialogues on AI have also been included in some regional trade agreements (RTAs) and digital economy agreements (see Chapter 3(b)(iii)).

Approaches to AI governance initiated at the regional level take different forms. Various regional initiatives have emerged in Africa, Asia and Latin America. Some of these take the form of ministerial declarations, such as the November 2023 Southern Common Market (MERCOSUR) Ministerial Declaration on the principles of human rights in the field of artificial intelligence,⁷⁰ the October 2023 Santiago Declaration to Promote Ethical Artificial Intelligence in Latin America and the Caribbean,⁷¹ and the May 2018 Declaration on AI in the Nordic-Baltic Region by the Nordic Council of Ministers.⁷² Others take the form of guides, such as the February 2024 Association of Southeast Asian Nations (ASEAN) Guide on AI Governance and Ethics (ASEAN, 2024), or strategy documents, such as the 2024 African Union Development Agency-New Partnership for Africa’s Development (AUDA-NEPAD) White Paper,⁷³ which led to the adoption, on 17 June 2024, of the African Continental Artificial Intelligence Strategy.⁷⁴

Some regional initiatives prioritize human rights and ethics, while others focus on economic development and growth. The MERCOSUR Declaration strongly emphasizes human rights and transparency, stressing the importance of avoiding discrimination, and of privacy and the integrity of information for democracy and the preservation of culture, while the Santiago Declaration focuses on human rights and ethics. The ASEAN Guide encourages alignment

on AI governance and ethics standards based on seven guiding principles,⁷⁵ but does not list inclusive growth, sustainable development and well-being – as per Principle 1 of the OECD AI Principles⁷⁶ – as a key principle. The ASEAN Guide includes recommendations for both domestic and regional initiatives⁷⁷ that governments in the ASEAN region can take to ensure the responsible design, development, and deployment of AI systems. Meanwhile, the AUDA-NEPAD White Paper and the African Union Continental Artificial Intelligence Strategy focus mainly on harnessing the potential of AI for economic development and growth, while promoting ethical use, minimizing potential risks and leveraging opportunities. The white paper stresses the importance of promoting innovation and building African multilingual tools through AI to support a “pan-African renaissance with AI” and lists five pillars of action: human capital development for AI, infrastructure and data, enabling environments for AI development and deployment, AI economy and encouraging investment in AI, and building sustainable partnerships. The Continental Strategy, adopted in June 2024, identifies four priority sectors: agriculture, healthcare, education and climate change adaptation. Likewise, the Arab AI Working Group focuses primarily on cooperation to reduce the digital divide and encourage capacity-building.

(iii) Regional trade agreements and digital economy agreements

AI-specific provisions have started to be incorporated into regional trade agreements (RTAs) and digital economy agreements,⁷⁸ but mainly take the form of soft – i.e. non-binding – provisions. While their incorporation into these agreements is positive, such provisions will not be sufficient to prevent regulatory fragmentation. Six agreements include AI-specific provisions. These are the United Kingdom-Australia Free Trade Agreement, the United Kingdom-New Zealand Free Trade Agreement, and the recently signed digital economy agreements between Australia and Singapore (SADEA), between Chile, New Zealand and Singapore (DEPA), between Singapore and the United Kingdom (UKSDEA), and between the Republic of Korea and Singapore (KSDPA), as well as a recently signed free trade agreement between Ukraine and the United Kingdom, which has not yet come into force.

AI provisions essentially take the form of best-effort clauses (i.e., which require parties to do everything possible to achieve the desired result). AI-specific provisions typically recognize the increasing importance of AI within the global economy and include best-effort clauses to either “collaborate and promote”⁷⁹ the development of governance frameworks to promote trusted, safe and responsible use of AI or “to develop”⁸⁰ such frameworks taking into account international guidelines, with the UK-Australia and UK-New Zealand agreements specifically referring to the 2019 OECD Principles (OECD, 2019a) (see Chapter 3(b)(iv)). Some agreements,

in particular those signed by the United Kingdom, also recognize the importance of a risk-based and outcome-based approach and of the principles of technological interoperability and technological neutrality,⁸¹ and include various cooperation provisions on exchanging information and sharing experiences and good practices on laws, regulations, policies, enforcement and compliance;⁸² ethical use, human diversity and unintended biases, industry-led technical standards and algorithmic transparency;⁸³ research;⁸⁴ and playing an active role in international fora,⁸⁵ with the UK-Australia and UK-Ukraine agreements explicitly referring to cooperation in the development of international standards, regulations and conformity assessment procedures.

Several AI-specific provisions explicitly refer to trade. Three agreements – United Kingdom-Ukraine,⁸⁶ United Kingdom-Singapore⁸⁷ and United Kingdom-Australia⁸⁸ – explicitly recognize the role of AI in promoting competitiveness and facilitating international trade. The United Kingdom-Australia agreement also encourages activities aimed at facilitating and promoting trade in emerging technologies, and the agreements between the United Kingdom and Ukraine and between the United Kingdom and Singapore encourage active participation in international fora “on matters concerning the interaction between trade and emerging technologies”.

Digital trade provisions included in RTAs are also important for AI development and use. The number of RTAs with digital trade provisions has been growing steadily since the early 2000s. The first digital trade provision can be found in the 2000 Jordan-United States Free Trade Agreement. By the end of 2022, 116 RTAs – representing 33 per cent of all existing RTAs – had incorporated provisions related to digital trade (López-González et al., 2023). These provisions typically include provisions on data flows, data localization, protection of personal information and access to government data, which, as seen in previous sections, play an important role in determining access to data needed to train AI models. Provisions that ban measures mandating disclosure of source code, software and algorithms have also been included in a number of trade agreements, most notably agreements led by the United States. Such provisions typically aim to protect technology firms from government measures requiring trade secrets to be disclosed as a prerequisite for operating in certain industries (Jones et al., 2024). Access to source code can, however, be important to assess the trustworthiness of AI systems (see Chapter 3 (a)(iii)). In addition, prohibitions on disclosure of source code can impact technology access and market competition, and limit the availability of open-source software (Jones et al., 2024). Provisions on source code can, therefore, have a significant impact on the development and use of AI and on promoting AI trustworthiness. Other provisions related to the adoption of standards and conformity assessment can also play a critical role in promoting trustworthy AI (see Chapter 3(a)(iii)), while provisions on competition in the digital market are important to address the market concentration power of AI (see Chapter 3(a)(i)). Finally, provisions on customs duties on electronic transmissions have been important in fostering an environment conducive to digital trade (IMF-OECD-UN-WBG-WTO, 2023).

The depth of digital trade provisions included in RTAs varies significantly, reflecting diverging approaches.

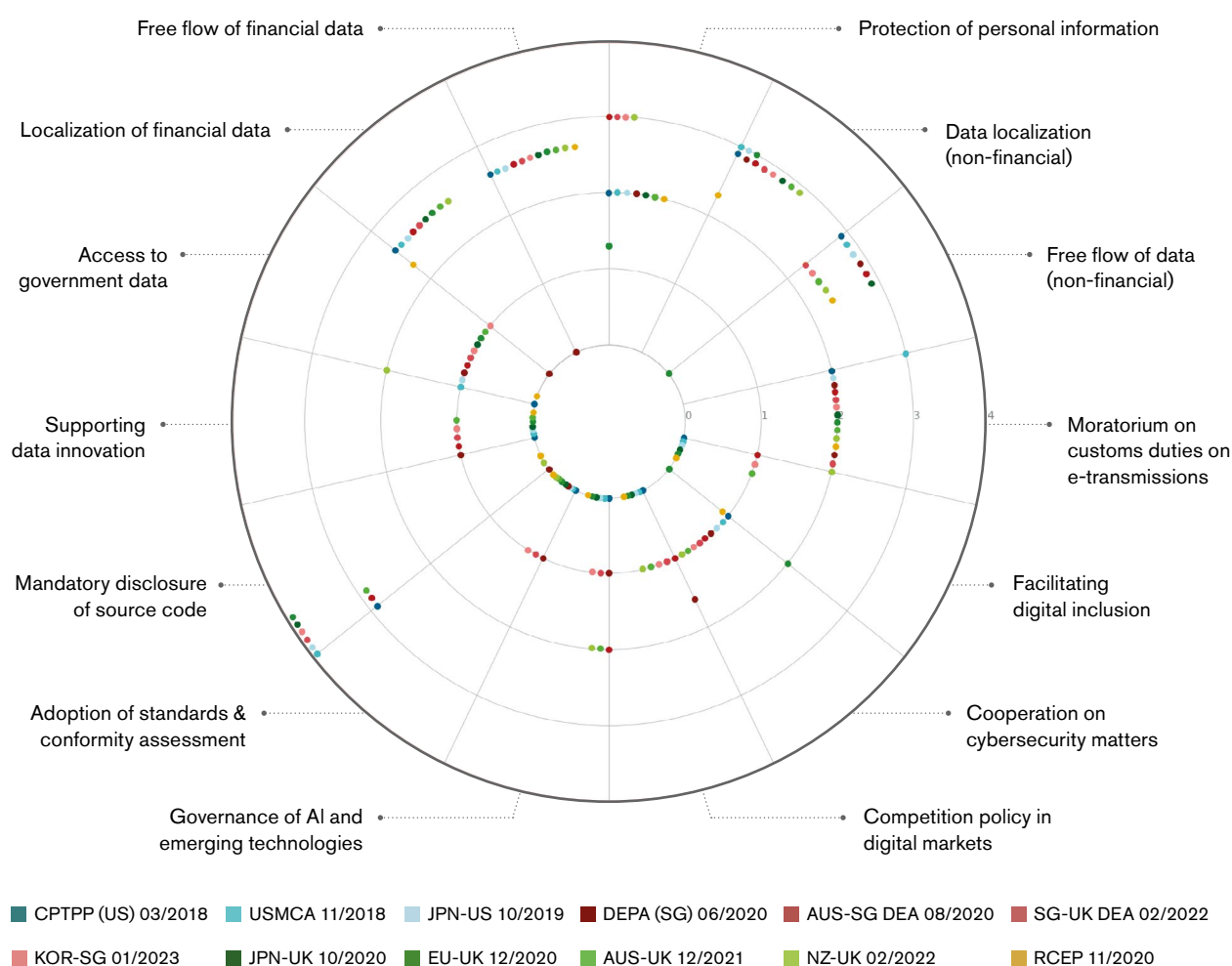
Analysing the digital trade provisions of 12 agreements concluded between March 2018 and January 2023, Jones et al. (2024) find a high degree of heterogeneity between the agreements (see Figure 3.7). For example, while most agreements contain binding obligations on the free flow of data, the United Kingdom-European Union RTA does not contain any provision on non-financial data flows. Regarding personal data protection, agreements led by the United States consider voluntary undertakings by private companies as sufficient to safeguard personal data, which contrasts with the European Union's comprehensive approach to data protection under the EU General Data Protection Regulation. Language on open government data takes the form of best-endavour language in agreements led by the United States and digital economy agreements,

while the United Kingdom-New Zealand Free Trade Agreement includes binding but non-specific language. As for disclosure of source code, agreements led by the United States and digital economy agreements include extensive and binding protection of source code, although digital economy agreements do not mention algorithms. In contrast, agreements signed by New Zealand and the Regional Comprehensive Economic Partnership (RCEP) do not include such provisions.⁸⁹

Few developing economies and LDCs have negotiated digital trade provisions.

The inclusion of detailed digital trade provisions tends to be more common in RTAs negotiated by high-income and certain middle- to upper middle-income economies. Only a handful of LDCs have engaged in RTAs that contain provisions related to digital trade (IMF-OECD-UN-WBG-WTO, 2023).

Figure 3.7: The depth of digital trade provisions included in RTAs varies significantly



Source: Authors' visualization based on Jones et al. (2024).

Notes: The value "0" (inner circle) means that the agreement does not contain a provision on a given issue. The value "1" means that the agreement contains a provision couched in purely hortatory, or exhortatory, language. The value "2" means that the commitment made is binding but non-specific. The value "3" means that commitments are binding and specific, with actions to be taken (or not taken) described in clear and precise language. The value "4" means that the commitments are binding and specific, with obligations that are more extensive in scope and very detailed. Points located between two lines correspond to the higher value in terms of commitments, but with flexibilities. The greater the flexibilities, the closer the point is to the lower-value inner circle.

Disciplines on trade in services in RTAs are also an important channel through which governments' trade policies and trade obligations can affect the policy environment for AI. However, the level of commitments undertaken differs significantly across economies. Services RTAs provide significantly higher levels of market access and national treatment commitments than under the WTO General Agreement on Trade in Services (GATS) for different modes of supply and services sectors, including for digital and AI-related services. For example, in the context of computer services, all WTO members from Europe, the Middle East and North America have undertaken some market access commitments on data processing services under the GATS and/or RTAs, and most WTO members have done so in Latin America and the Caribbean (88 per cent) and in Asia (91 per cent). However, in Africa, 26 per cent of WTO members have market access commitments on data processing services, whether under the GATS or RTAs, although that proportion will increase when the services commitments of the African Continental Free Trade Area (AfCFTA) enter into force and are notified to the WTO (Roy and Sauvé, forthcoming).⁹⁰

(iv) International initiatives to address the challenges raised by AI

Policy initiatives

The last few years have witnessed a wave of various initiatives related to AI. This impetus has been driven by the realization that the inherently international nature of the risks and benefits associated with AI require discussion, cooperation and solutions that are also international in nature (see Figure 3.8).⁹¹

These initiatives involve different stakeholders and take different forms. International initiatives involve a broad range of stakeholders, including governments, intergovernmental organizations, international standard-setting bodies and businesses. Most of these initiatives take the form of high level principles, guidance, voluntary recommendations, scientific reports, codes of conduct, or lists of policy examples, while others take the form of international standards. In May 2024, the first binding treaty on AI – the Council of Europe Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law⁹² – was adopted.

There are elements of complementarity among such initiatives, and alignment on core principles, but different initiatives prioritize different aspects of AI governance. Some of the common themes across various international initiatives include, for instance, promoting safe, secure, trustworthy, “human-centric”, ethical, transparent, accountable and interoperable AI, and identifying and mitigating AI-triggered risks through various actions, domestic policies and international cooperation.⁹³ However, in certain instances, international initiatives appear to prioritize these themes differently. Some initiatives focus on issues like human

rights and the ethics of AI, such as the United Nations Educational, Scientific and Cultural Organization (UNESCO) Recommendation on the Ethics of AI, while others are centred around safety, security, the trustworthiness of AI or its interoperability, such as the Bletchley Declaration on AI Safety.

A number of initiatives also contain various common elements that have an important trade and WTO angle.

These include:

- the recognition of the role of regulations and standards (including certification procedures) in governing AI and the importance of interoperability between such tools;
- the need to avoid regulatory fragmentation by using international standards to govern AI;
- the importance of an appropriate and balanced approach to protecting and enforcing IP rights;
- the importance of privacy, personal data protection and data governance;
- the importance of international cooperation, coordination and dialogue.

Importantly, explicit references to the WTO were included in the Final Report of the UN AI Advisory Body. The Final Report stresses the need for “proper orchestration” and coordination among the many international processes and organizations producing key documents related to AI governance, to enable a “shared normative foundation for all AI-related efforts”, expressly referring to various WTO agreements, such as the General Agreement on Tariffs and Trade (GATT), the General Agreement on Trade in Services (GATS), the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), the Technical Barriers to Trade (TBT) Agreement, the Information Technology Agreement (ITA) and the Trade Facilitation Agreement (TFA) (paragraph 76 and Figure 9 of the Final Report). The Final Report also notes the pivotal role of international standards and regulatory cooperation and recognizes the key role of the WTO in this area (paragraph 121).⁹⁴ Finally, it recommends the creation of a “Global AI Data Framework” involving a variety of key actors, including economies and relevant international organizations, including the WTO (paragraph 170). More detail can be found in Annex 3.

Several of these initiatives also address the environmental impacts of AI. This is the case, for example, for the OECD AI Principles, the G20 AI Principles, the New Delhi Leaders' Declaration,⁹⁵ the UNESCO AI Recommendation, the G7 Guiding AI Principles and the G7 AI Code of Conduct, the Bletchley Declaration on AI Safety, and the UN Advisory Board on AI in the United Nations. Further, the International Organization for Standardization (ISO)/International Electrotechnical Commission (IEC) Joint Technical Committee (JTC) 1 and subcommittee 42 (SC) (ISO/IEC JTC 1/SC 42) is currently developing an international standard specifically about “environmental sustainability aspects of AI systems.” AI policy was a key issue at the G20 Summit in Rio de Janeiro in November 2024, with a focus on the use of AI for sustainable development; in the G7 Trieste Ministerial Declaration adopted in 2024, G7 economies also expressed their desire to participate in the G20 AI and sustainability discussions.

However, there is still no global alignment on AI terminology. Global agreement over key AI terminology and definitions may be a particularly important trade-related element, as it may help to ensure coherence and interoperability and to avoid fragmentation across various domestic AI regulatory regimes (Meltzer, 2023). As explained in this report, regulatory fragmentation can itself represent an important trade barrier, in particular for developing economies and micro, small and medium-sized enterprises (MSMEs). In this respect, the OECD AI Principles⁹⁶ contain various AI definitions, of which the definitions of an “AI system”⁹⁷ and an “AI system lifecycle”⁹⁸ are key for the implementation of any domestic AI strategy or policy and, in particular, for regulation. The Council of Europe Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law⁹⁹ also contains a definition of an “AI system” which is virtually identical to that in the OECD Principles.¹⁰⁰ The ISO/IEC JTC 1/SG 42, which is dedicated to AI standard-setting, adopted in 2022 a document¹⁰¹ containing a wide range of detailed definitions and terminology in the field of AI. It included a definition of an “AI system”, which shares some similarities but also includes some differences with the definition in the OECD Principles. Finally, while the G20 AI Principles¹⁰² have more or less integrated all of the OECD Principles, they do not expressly endorse the definitions, including that of an “AI System”. Unlike OECD and ISO/IEC, the UNESCO Recommendation on AI Ethics does not define AI.¹⁰³

Some initiatives seem to be moving beyond general principles or guidance into implementing more targeted or specific actions. For instance, in order to foster their knowledge on existing approaches and practices, the G20 launched the “Examples of National Policies to Advance the G20 AI Principles”,¹⁰⁴ and the G20 “Policy Examples on How to Enhance the Adoption of AI by MSMEs and Start-up”.¹⁰⁵ In 2024, the G7 announced plans to advance its 2023 Hiroshima AI process. The planned actions include expanding outreach to partner governments to broaden support for the G7 AI Guiding Principles and Code of Conduct, intensifying efforts to encourage adherence to these two instruments, and intensifying cooperation across multilateral forums to promote the G7 vision for advanced AI systems.¹⁰⁶ In addition, following up on the 2023 Bletchley Declaration on AI Safety, governments have agreed to convey a panel of experts to produce an Intergovernmental Panel on Climate Change (IPPC)-like “State of the Science” Report,¹⁰⁷ which will aim to review the latest cutting-edge research on the risks and capabilities of frontier AI models. The interim *International Scientific Report on the Safety of Advanced AI* was published in May 2024¹⁰⁸ and summarizes the best of existing research, while identifying areas of research priority. It does not make policy or regulatory recommendations, but instead aims to inform both domestic and international policymaking. The final report is expected to be published ahead of the next AI summit which is expected to be held in February 2025 in France (see also Annex 3).

The significant overlap between initiatives, the differing priorities and the lack of agreement on key terminology could create implementation challenges. This may limit efforts to prevent fragmentation. Alignment

on core principles does not guarantee alignment on how such principles can be implemented in practice. In the absence of strong coordination, current international initiatives may not be sufficient to prevent regulatory fragmentation at the global level. The need to improve coordination was acknowledged in the Final Report (2024) of the UN AI Advisory Body (AIAB) published in September 2024 and the Global Digital Compact adopted by the UN General Assembly in September 2024. The Final Report identifies three “global AI governance gaps” to be addressed: a “representation” gap, a “coordination” gap and an “implementation” gap. The WTO is relevant for all three, and as noted above, specific references to the WTO are included in various places of the Final Report. As for the Global Digital Compact, it includes a commitment by UN members to initiate a Global Dialogue on AI governance involving governments and all relevant stakeholders (paragraph 56).

International initiatives to close the AI divide

Increasingly, international organizations are developing courses on AI and are integrating AI in their technical assistance activities, some of which have a trade component. The International Telecommunication Union (ITU), for example, offers an online course titled “The governance of artificial intelligence” and, in partnership with 40 other UN agencies, the ITU launched “AI for Good,” an action-oriented global platform on AI to identify practical applications of AI to advance the UN Sustainable Development Goals (SDGs).¹⁰⁹ AI for Good includes a year-round online programme of webinars, with an annual in-person AI for Good Global Summit. Other specialized UN agencies have developed projects focused on their own areas of expertise. UNESCO, for example, has developed a Readiness Assessment Methodology to support its members in their implementation of the UNESCO Recommendation on the Ethics of AI, and is providing targeted technical assistance in this context through projects such as its “AI needs assessment in African countries” programme.¹¹⁰ Meanwhile, the United Nations Industrial Development Organization (UNIDO) has been organizing dialogues on “Empowering SMEs in Developing Countries through Artificial Intelligence”¹¹¹ to promote AI adoption by MSMEs in developing economies, to enhance their competitiveness and sustainability through shared conversations. A related publication by UNIDO includes practical recommendations and tools to help MSMEs navigate challenges and leverage AI for various business functions and production areas. As for the World Bank, two notable projects with an AI dimension are the “Machine learning in Algeria” project, which aims to enhance efficiency and integrity in customs operations using machine learning, and “Fraud analytics in Kenya using AI applications”, which aims to improve revenue collection through anti-fraud measures.¹¹² And the United Nations Interregional Crime and Justice Research Institute (UNICRI) has developed a course for law enforcement agencies to equip them with the necessary resources to institutionalize responsible AI, ensuring its alignment with human rights and ethics.¹¹³

Figure 3.8: Key international policy initiatives in the area of AI



* Signatories include Australia, Canada, the European Union, France, Germany, Italy, Japan, the Republic of Korea, Singapore, the United Kingdom and the United States.

The UN AI Advisory Body (AIAB) has called for the establishment of a global fund for AI. Published in September 2024, the Final Report of the UN AIAB recommends the creation of a global fund for AI to “put a floor under the AI divide”. Managed by an independent governance structure, the fund would receive financial and in-kind contributions from public and private sources to facilitate access to AI enablers, such as shared computing

resources for model training, sandboxes and curated datasets “to catalyse local empowerment for the SDGs”,¹¹⁴ The Global Digital Compact,¹¹⁵ which was adopted by the United Nations General Assembly in September 2024 after the publication of the UN AIAB report, calls for “innovative voluntary financing options for artificial intelligence capacity-building that take into account the recommendations of the High-level Advisory Body on Artificial Intelligence on a Global Fund on AI”.

Endnotes

1 See <https://ourworldindata.org/extreme-poverty-in-brief>.

2 See <https://eto.tech/>.

3 See <https://eto.tech/>.

4 See <https://oecd.ai/en/data>.

5 See <https://stackoverflow.com/>.

6 See <https://ourworldindata.org/grapher/share-new-artificial-intelligence-cs-phds-female>.

7 At the same time, AI also holds procompetitive potential. For instance, it empowers consumers to utilize abundant data for personalized products and transactions, and guides them in navigating complex or uncertain markets to select the best offers based on preferences. This may lead to the emergence of “algorithmic consumers”, whose decision-making is partially automated through algorithms (Gal and Elkin-Koren, 2017).

8 Various competition enforcement cases were recently launched against AI companies. For example, the US Federal Trade Commission (FTC) went to court to block a proposed acquisition of Arm Ltd. by Nvidia, one of the leading producers of advanced chips powering AI, which resulted in the latter abandoning the deal (see <https://www.ftc.gov/news-events/news/press-releases/2022/02/statement-regarding-termination-nvidia-corporation-attempted-acquisition-arm-ltd>). The European Commission, like the UK Competition Markets Authority and the FTC, also started looking into whether the investment of Microsoft in OpenAI constituted a merger (European Commission, 2024a). Cognizant of the risks that AI poses for competition, the competition authorities of the European Union, the United Kingdom and the United States of America issued in July 2024 a Joint Statement on Competition in Generative AI Foundation Models and AI Products laying out various principles for protecting competition in the AI ecosystem (see <https://www.gov.uk/government/publications/joint-statement-on-competition-in-generative-ai-foundation-models-and-ai-products/joint-statement-on-competition-in-generative-ai-foundation-models-and-ai-products>).

9 An example of this is the fact that the Google search engine can outperform that of Microsoft because the former has wider access to rarer queries. Having a variety of data and, in particular, its ability to capture more rare events are also important for making better predictions (Goldfarb and Trefler, 2018).

10 However, advancements such as federated learning, which allows entities in various locations to build machine learning models collaboratively, without exchanging data (it is the algorithm that is transferred, not the data itself), and data trusts, a system and legal entity that manages someone's data on their behalf, could mitigate the challenges linked to cross-border data flows (Bonawitz et al., 2019; World Economic Forum, 2020a).

11 AI can turn even non-personal enterprise and operational data, such as stock inventory, into privacy risks. These data,

stored within corporate networks, can inadvertently reveal information about personnel involved in data collection or analysis. In addition, metadata in online communications, such as phone numbers, emails or IP addresses, can make users identifiable even if the data do not directly reveal personal identities (Lee-Makiyama, 2018).

12 See <https://www.consilium.europa.eu/en/press/press-releases/2024/10/10/eu-brings-product-liability-rules-in-line-with-digital-age-and-circular-economy/#:~:text=The%20EU's%20product%20liability%20regime,caused%20the%20injury%20or%20damage> and <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022PC0496>.

13 Trustworthiness is mentioned in international AI principles and declarations as a key attribute that an AI system should possess. See, for example, <https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023>, https://www.mofa.go.jp/policy/economy/g20_summit/osaka19/pdf/documents/en/annex_08.pdf, Organisation for Economic Co-operation and Development (2019a) and United Nations Educational, Scientific and Cultural Organization (2021). In AI terminology, “trustworthiness” means the “ability to meet stakeholder ... expectations in a verifiable way” (e.g., via certification against technical specification in a regulation or standard). More specifically, the trustworthiness of an AI system relates to its ability to meet various expectations, for example in terms of its “reliability”, “availability”, “resilience”, “security”, “privacy”, “safety”, “accountability”, “transparency”, “integrity”, “authenticity”, “quality” and “usability”. See ISO/IEC standard 22989:2022, sub clause 3.5.16 (Trustworthiness - definition) and clause 5.15 (Trustworthiness - concept). See also ISO/IEC TR 24028:2020. While the composite term “safe and trustworthy” AI is frequently used, given that “safety” is subsumed into the above definition of trustworthiness, in this report, for simplicity, we will only refer to trustworthy AI.

14 For instance, certain risks may be associated with AI-enabled autonomous vehicles that stem not from the physical components of the vehicle. Instead, the AI algorithm (and how it has been trained), may lead the vehicle to “behave” in a risky manner, causing not only material harms (e.g., physical injuries to the driver, passengers or pedestrians) but also, uniquely, immaterial harms (e.g., privacy, cybersecurity, etc.). See UK Parliament House of Commons' Report on Self-Driving Vehicles (HC 519, 15 Sep 2023), paragraph 66 (noting studies warning that “fleets or models of self-driving vehicles could be targeted by ‘malicious, possibly terrorist, systemic hacking’”). Regulatory solutions to such immaterial risks may also present complex ethical questions, e.g., the famous “trolley problem”, whereby an autonomous vehicle has to “choose”, for example, between colliding with an elderly person and colliding with a mother and her young child. (e.g., Wells (2023); Lin (2021)).

15 “Injury to pedestrians due to the malfunction of an autonomous vehicle AI system would be tangible physical harm. Some harms, however, such as psychological harms, may not be as tangible or quantifiable. Other aspects of harm that may be intangible or difficult to directly observe include bias or discrimination that may disproportionately and negatively impact particular communities but be difficult to observe at the level of the individual. Violations of the fundamental right to privacy may also be intangible, such as the non-transparent use of an employee monitoring AI system”. OECD Working Party on Artificial Intelligence Governance: Stocktaking for the development of an AI incident definition, document EP/AIGO(2022)11/FINAL (21 Oct 2023).

16 See, e.g., Report from the European Commission to the European Parliament, the Council and the European Economic and Social Committee on “The Safety And Liability Implications of Artificial Intelligence, the Internet of Things and Robotics”, COM(2020) 64 final (19 February 2020), page 8. See <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064>.

17 See <https://www.iso.org/committee/6794475/x/catalogue/p/1/u/0/w/0/d/0>.

18 The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement) addresses inter alia the protection of trade secrets, including imposing certain conditions when proprietary information (“undisclosed test and other data”) is accessed and used by governments for regulatory purposes, albeit only in the context of “marketing approval” (e.g., conformity assessment procedures such as product certification and approval) of pharmaceuticals and agricultural chemical products (Article 39.3). Similarly, the Technical Barriers to Trade (TBT) Agreement requires that WTO members ensure that the confidentiality of information in the context of conformity assessment procedures (e.g. product certification and approval) is (i) respected for imported and national products “in the same way” and (ii) respected “in such a manner that legitimate commercial interests are protected.” (Article 5.2.4).

19 Mitchell et al. (2023) contains a detailed analysis of circumstances when regulating AI can be performed without need to access source code (“white box” testing for low-risk AI systems), and of circumstances when a deeper understanding and explanation of the AI system’s decisions and recommendations is needed (high risk AI systems) and justifies requiring access to the code (“black box” testing).

20 “Evolution” in the sense that some AI systems allow the product to better perform, adapt and finetune overtime for a given circumstance or for a given user, as it works in practice and receives and crunches more data; a sort of “personalized AI product” similar to the idea of “personalized medicine” (e.g., using knowledge of a patient’s genetic profile to select “the proper medication or therapy and administer it using the proper dose or regimen” – see <https://www.genome.gov/genetics-glossary/Personalized-Medicine>). In fact, AI can be a driver and enabler for advancing personalized medicine in the area of genomics medicine. See Cesario et al. (2023); World Health Organization (2021).

21 In this respect, the EU AI Act (2024a), for instance, notes in its preamble (recital 5), that “AI may generate risks and cause harm to public interests and fundamental rights” and that “[s]uch harm might be *material or immaterial*, including physical, psychological, societal or economic harm.” (italics added).

22 Lund et al. (2023) present a useful example of this tension with respect to medical devices, explaining that one of the “main obstacles of using AI in healthcare, and therefore AI-based medical software”, is that “how to address continuous change i.e., locked algorithms vs non-locked autonomous systems is a challenge.

The strength and advantages with AI/ML are the ability to train and improve the system based on new real-world data. However, the system also needs to be continuously safe for patients and other users, as well as comply with the applicable regulations regarding, for example, validation”.

23 For instance, “[c]ustomisation makes traceability and enforcement of product safety and cybersecurity more challenging – many products (or properties) are changing constantly” (Lund et al. 2023).

24 The EU AI Act (2024a), for instance, seems to contain certain provisions on this issue, as it requires that AI systems be re certified if, after deployed, they present unforeseen “substantial modifications” (as defined in Article 3(23)), i.e., “... whenever a change occurs which may affect the compliance of a high-risk AI system with this Regulation (e.g. change of operating system or software architecture), or when the intended purpose of the system changes, that AI system should be considered a new AI system which should undergo a new conformity assessment”. However, “changes occurring to the algorithm and the performance of AI systems which continue to ‘learn’ after being placed on the market or put into service (i.e., automatically adapting how functions are carried out) should not constitute a substantial modification, provided that those changes have been pre-determined by the provider and assessed at the moment of the conformity assessment”. The EU AI Act also foresees that ex post marketing surveillance over AI products may: “ensure that the possible risks emerging from AI systems which continue to ‘learn’ after being placed on the market or put into service can be more efficiently and timely addressed”.

25 ISO/IEC TR 24368 (2022) gives examples of areas in which there is an “increasing risk for undesirable ethical and societal outcomes and harms”, e.g.: “financial”; “psychological”; “physical health or safety”; “intangible property (for example, IP theft, damage to a company’s reputation)”; “social or political systems (for example, election interference, loss of trust in authorities)”; and “civil liberties (for example, unjustified imprisonment or other punishment, censorship, privacy breaches)”.

26 Commenting on the fact that the AI Act’s implementation may involve the adoption of technical standards for addressing both material (e.g., health) and immaterial risks (e.g., fundamental rights), Smuha and Yeung (2024), observe that: “... unlike risks to safety generated by chemicals, machinery or industrial waste, all of which can be materially observed and measured, fundamental rights are, in effect, political constructs. These rights are accorded special legal protection so that an evaluation of alleged interference requires close attention to the nature and scope of the relevant right and the specific, localized context in which a particular right is allegedly infringed. We therefore seriously doubt whether fundamental rights can ever be translated into generalized technical standards that can be precisely measured in quantitative terms, and in a manner that faithfully reflects what they are, and how they have been interpreted under the European Charter on Fundamental Rights and the European Convention on Human Rights”.

27 See also WTO official document number G/TBT/GEN/356, available at <https://docs.wto.org/>.

28 See, for example, the European Union’s Artificial Intelligence Act (AIA) (<https://artificialintelligenceact.eu/>), under which all providers of “general-purpose AI models” are mandated to create technical documentation that details the training and testing processes, establish a copyright policy, and provide a sufficiently detailed summary of the content used for training. In contrast, free and open AI models are only obliged to establish a copyright policy and submit a summary of training content.

29 RDC n° 657/2022 – Anvisa https://antigo.anvisa.gov.br/documents/10181/5141677/RDC_657_2022_.pdf/f1c32f0e-21c7-415b-8b5d-06f4c539bbc3.

30 Such as Israel, Japan, the Republic of Korea, Singapore, the United Kingdom and the United States. The Global AI Index 2024 published by Tortoise Media, which ranks economies by their AI capacity at the international level: <https://www.tortoisemedia.com/intelligence/global-ai/>.

31 While copyright protects creation of the (human) mind, patent protection is available for technical innovations (by humans).

32 US Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, available at https://copyright.gov/ai/ai_policy_guidance.pdf; US Court of Appeals for the 9th Circuit, *Naruto v. Slater*, <https://cdn.ca9.uscourts.gov/datastore/opinions/2018/04/23/16-15469.pdf>; US District Court for the District of Columbia, *Thaler v. Perlmutter*, https://ecf.dcd.uscourts.gov/cgi-bin/show_public_doc?2022cv1564-24.

33 CJEU, *Infopaq International A/S v Danske Dagblades Forening*, Case C-5/08 (Intellectual Property Repository, 2023; Zhou, 2019).

34 Strategies articulate the government's vision regarding the contribution of science, technology and innovation (STI) to the social and economic development of an economy. They set priorities for public investment in STI and identify the focus of government reforms, for instance in areas such as funding public research and promoting business innovation (OECD, 2016b).

35 See OECD AI database. <https://oecd.ai/en/dashboards/overview>

36 Although not a country, the European Union has the power to adopt EU-wide trade-related legislation within the parameters set by its founding treaties.

37 See https://ec.europa.eu/commission/presscorner/detail/en/ip_24_383.

38 See <https://digital-strategy.ec.europa.eu/en/policies/plan-ai>.

39 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). AIA was published in the EU Official Journal on 12 July 2024 and entered into force 20 days later. However, most of AIA's rules are only applicable 24 months after its entry into force, although it provides for shorter applicability periods with respect to certain rules (e.g., bans on “prohibited practices” that are listed as posing “unacceptable risks” will already apply six months after entry into force), as well as longer periods for others (e.g., 36 months for certain “high risk systems” covered by existing EU harmonization legislation and for general-purpose AI systems on the EU market before the Act applies to them).

40 In the notification in 2021 of a draft of the AIA (document G/TBT/N/EU/850), the European Union explained that this proposal was meant to provide: “... a set of recommendations intended to help the organization develop, provide, or use AI systems responsibly in pursuing its objectives and meet applicable requirements, obligations related to interested parties and expectations from them. It includes the following: approaches to establish trust in AI systems through transparency, explainability, controllability, etc.; engineering pitfalls and typical associated threats and risks to AI systems, along with possible mitigation techniques and methods; and approaches to assess and achieve availability, resiliency, reliability, accuracy, safety, security and privacy of AI systems. This document is applicable to any

organization, regardless of size, type and nature, that provides or uses products or services that utilize AI systems”.

41 However, some experts argue that some important provisions of the AI Act do not follow a purely risk based approach (Ebers, 2024).

42 See <https://oecd.ai/en/>.

43 For example, the UK has committed £100 million toward building a “public foundation model” to support academic, small business and public sector applications (see <https://www.gov.uk/government/news/initial-100-million-for-expert-taskforce-to-help-uk-build-and-adopt-next-generation-of-safe-ai>), and the US National Artificial Intelligence Research Resource is working in a similar direction (see <https://new.nsf.gov/focus-areas/artificial-intelligence/nairr>).

44 Standards are one of the three types of technical barriers to trade (TBT) measures that establish product specifications. They differ from technical regulations, however, as standards are voluntary documents. It is also not uncommon for standards adopted by governments to be made mandatory later on, thus becoming technical regulations. Standards can be developed by different entities within WTO Members, including both governmental and non-governmental bodies (WTO, 2021).

45 See, e.g., Kerry (2024). For instance, in 2024, China issued draft Guidelines for AI Standardisation which proposes to form more than 50 national and industry-wide standards and more than 20 international standards for AI by 2026 (see <https://www.reuters.com/technology/china-issues-draft-guidelines-standardising-ai-industry-2024-01-17/> and <https://mmlcgroup.com/miit-ai/>). The European Commission mandated European standardisation organizations to develop AI-related standards taking into account that standards will play an important role in fulfilling requirements under the EU AI Act ([https://ec.europa.eu/transparency/documents-register/detail?ref=C\(2023\)3215&lang=en](https://ec.europa.eu/transparency/documents-register/detail?ref=C(2023)3215&lang=en)).

46 For example, China's standards strategy of 2021 identified AI as one of the key areas. See Kerry (2024).

47 This is based on the data from the AI Standards Hub and is provided for illustration purposes only. This data is presented without prejudice, and should not be understood as a position, on whether these documents are “standards” within the definition of Annex 1 of the TBT Agreement. See AI Standards Search - AI Standards Hub.

48 See the AI Standards Hub at <https://aistandardshub.org>.

49 For example, the WTO Technical Barriers to Trade (TBT) Agreement requires that WTO members use relevant international standards as a basis of their domestic standards, technical regulations and certification procedures (see Chapter 4).

50 See <https://webarchive.nla.gov.au/awa/20220816053410/https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-action-plan>.

51 See http://gd.china-embassy.gov.cn/eng/zxhd_1/202310/t20231024_11167412.htm.

52 See <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>. See also NIST “A Plan for Global Engagement on AI Standards” (final, July 2024 - available at: <https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-5.pdf>) and US Government National Standards Strategy for Critical and Emerging Technology presented at the TBT Committee meeting held on 21-23 June 2023 (<https://docs.wto.org/dol2fe/Pages/SS/directdoc.aspx?filename=q:/G/TBT/M90.pdf&Open=True>, paragraph 6.32).

53 For example, Australia has been actively engaged in the work of the International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) Joint Technical Committee (ISO/IEC JTC1/SC42) and, in 2024, Australia announced the adoption of one of the ISO/IEC JTC1/SC42 standards (see <https://www.standards.org.au/news/standards-australia-adopts-the-international-standard-for-ai-management-system-as-iso-iec-42001-2023>).

54 EU AI Act, Preamble, Recital (176).

55 While there is no globally agreed definition, the European Parliament Research Service notes in its paper on “Artificial Intelligence Act and Regulatory Sandboxes” ([https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI\(2022\)733544_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf)) that “regulatory sandboxes generally refer to regulatory tools allowing businesses to test and experiment with new and innovative products, services or businesses under supervision of a regulator for a limited period of time. As such, regulatory sandboxes have a double role: 1) they foster business learning, i.e. the development and testing of innovations in a real-world environment; and 2) support regulatory learning, i.e. the formulation of experimental legal regimes to guide and support businesses in their innovation activities under the supervision of a regulatory authority”.

56 See <https://aiverifyfoundation.sg>.

57 i.e., the initial version of a product that includes only the core features necessary to meet basic user needs and gather feedback for future improvements.

58 WTO official documents IP/N/1/JPN/36, IP/N/1/JPN/C/6 and IP/C/M/92/Add.1, available at <https://docs.wto.org/>.

59 The UK Copyright, Designs and Patents Act 1998 provides that authorship is attributed to “the person by whom the arrangements necessary for the creation of the work are undertaken”, paragraph 9(3). Other common law jurisdictions such as India (copyright Act 1957 paragraph 2(d)), Ireland (Copyright and Related Rights Act 2000 21), New Zealand (Copyright Act 1994 5(1)) and South Africa (Copyright Act 1978 1(iv)) follow the UK approach.

60 For more information on developing economies pursuing open government data policies see Verhulst and Young (2017).

61 See <https://data.gov.in/>.

62 See <https://perma.cc/JAX3-55U8>. The OECD recently launched an Open Government Data project to map practices across economies and assess the impact of open government data (OECD, 2019b).

63 Unlike anti-trust policies, ex ante regulations apply at an industry or sectoral level and attempt to define how the largest companies must compete in the market. One such set of regulations is the European Union Digital Markets Act (DMA), which entered into force in November 2022 and became applicable, for the most part, on 2 May 2023. The DMA is designed to address the market power of major digital platforms, referred to as “gatekeepers”. It aims to ensure fair competition and innovation in the digital market by preventing gatekeepers from imposing unfair conditions on businesses and consumers (European Commission, 2022). The DMA includes specific obligations for these gatekeepers, such as allowing third parties to interoperate with their services and prohibiting them from favouring their own services. The UK Digital Markets, Competition and Consumers Bill is another example of new ex ante approach to digital markets. The Bill encourages the most powerful firms in dynamic digital markets to work with regulators to ensure that competition is maintained on an ongoing basis. See <https://www.gov.uk/government/news/changes-to-digital-markets-bill-introduced-to-ensure-fairer-competition-in-tech-industry#:~:text=The%20>

Digital%20Markets%2C%20Competition%20and,in%20and%20innovate%20new%20technology.

64 The Safe Harbour Privacy Principles, which were developed between 1998 and 2000 to prevent private organizations within the European Union or United States that store customer data from accidentally disclosing or losing personal information, were brought down by the European Court of Justice (ECJ) in 2020 after Max Schrems, an Austrian activist, lawyer and author brought a case against Facebook for its privacy violations, including violations of European privacy laws and the alleged transfer of personal data to the US National Security Agency (NSA) as part of the NSA’s PRISM data-mining programme. The Safe Harbour Privacy Principles were replaced with the Privacy Shield until 2020, when the ECJ once again brought it down. A new agreement was reached in July 2023 to allow data flows based on the “adequacy decision” mechanism of the EU General Data Protection Regulation.

65 The United States initiated export controls on semi-conductors in 2022, and these restrictions were broadened over time. In 2023, the Netherlands imposed restrictions on high end chipmaking. The United Kingdom, Canada and Japan followed with their own restrictions. See Financial Times (2022) and Wolff (2022).

66 See https://www.wto.org/english/tratop_e/dispu_e/cases_e/ds615_e.htm.

67 See <https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence>.

68 See <https://digital-strategy.ec.europa.eu/en/library/ai-public-good-eu-us-research-alliance-ai-public-good>.

69 See <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/12/u-s-singapore-critical-and-emerging-technology-dialogue-joint-vision-statement/>.

70 See <https://www.raadh.mercosur.int/wp-content/uploads/2024/04/DECLARACION-SOBRE-LOS-PRINCIPIOS-DE-DERECOS-HUMANOS-EN-EL-AMBITO-DE-LA-INTELIGENCIA-ARTIFICIAL.pdf>.

71 See https://minciencia.gob.cl/uploads/filer_public/40/2a/402a35a0-1222-4dab-b090-5c81bbf34237/declaracion_de_santiago.pdf.

72 See <https://www.stjornarradid.is/library/04-Raduneytin/ForsAetisraduneytid/Framtidarnefnd/AI%20in%20the%20Nordic-Baltic%20region.pdf>.

73 See <https://dig.watch/resource/auda-nepad-white-paper-regulation-and-responsible-adoption-of-ai-in-africa-towards-achievement-of-au-agenda-2063>.

74 See <https://au.int/en/pressreleases/20240617/african-ministers-adopt-landmark-continental-artificial-intelligence-strategy#:~:text=The%20Continental%20AI%20Strategy%20provides,potential%20risks%2C%20and%20leveraging%20opportunities>.

75 The seven guiding principles are transparency and explainability, fairness and equity, security and safety, robustness and reliability, human-centricity, privacy and data governance, and accountability and integrity.

76 See <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>.

77 National recommendations include nurturing AI talent and upskilling the workforce, supporting the AI innovation ecosystem and promoting investment in AI start-ups, investing in AI research and development, promoting adoption of useful tools by businesses to implement the ASEAN Guide on AI Governance and Ethics, and raising awareness among citizens on

the effects of AI in society. The regional recommendations are: to establish an ASEAN Working Group on AI Governance consisting of representatives from member states to drive and oversee AI governance initiatives in the region; to adapt the AI Guide to address the governance of generative AI; and to compile a compendium of use cases demonstrating practical implementation of the AI Guide by organizations operating in ASEAN.

78 Digital economy agreements are a new type of agreement. They aim to regulate digital trade, data flows and emerging technologies like AI. Digital economy agreements reflect governments' response to the need for regulatory frameworks tailored to the complexities of digital trade and the digital economy. To date, four digital economy agreements have been signed and have entered into force: the Singapore-Australia Digital Economy Agreement (SADEA), signed in 2020; the Digital Economy Partnership Agreement (DEPA) between Chile, New Zealand and Singapore, signed in 2020; the United Kingdom-Singapore Digital Economy Agreement (UKSDEA), signed in 2022; and the Republic of Korea-Singapore Digital Partnership Agreement (KSDPA), signed in 2022. Others under negotiation include the ASEAN Digital Economy Framework Agreement (DEFA) and the EFTA-Singapore Digital Economy Agreement.

79 KSDPA, DEPA and United Kingdom-Australia.

80 United Kingdom-Australia, United Kingdom-New Zealand and United Kingdom-Singapore (the latter specifies "where appropriate").

81 United Kingdom-Ukraine, United Kingdom-Singapore and United Kingdom-New Zealand.

82 United Kingdom-Ukraine, KSDPA, United Kingdom-Singapore and United Kingdom-New Zealand.

83 United Kingdom-Ukraine, United Kingdom-Singapore and United Kingdom-New Zealand.

84 United Kingdom-Ukraine, United Kingdom-Singapore, SADEA and United Kingdom-Australia.

85 United Kingdom-Ukraine and United Kingdom-New Zealand.

86 Article 132-V.

87 Article 8.61-R.

88 Article 20.4.

89 New Zealand decided to exclude provisions on source code from its agreements following a November 2021 decision of the Waitangi Tribunal, which found the source code provision in the Comprehensive and Progressive Agreement for Trans-Pacific Partnership (CPTPP) to be in breach of the Treaty of Waitangi after Māori tech experts argued that there was a risk of biased assumptions in algorithmic design or training data. See Jones (2024).

90 The analysis in Roy and Sauvé (forthcoming) is based on 142 RTAs notified under GATS Article V.

91 See, for example, the Bletchley Declaration (2023b), which states that, "[m]any risks arising from AI are inherently international in nature, and so are best addressed through international cooperation".

92 See <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.

93 While, so far, AI global governance initiatives have, on the one hand, "yielded similarities in language, such as the importance of fairness, accountability, and transparency", on the other hand, approaches on defining AI are less coordinated and coherent (https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf).

94 i.e., calling for a "AI Standards Summit" involving key international standard-setting bodies (e.g., International Telecommunication Union (ITU), the International Organization for Standardization (ISO) / International Electrotechnical Commission (IEC) and the Institute of Electrical and Electronics Engineers (IEEE)) and expressly indicating that the WTO should be involved in these discussions.

95 See <https://www.mea.gov.in/Images/CPV/G20-New-Delhi-Leaders-Declaration.pdf>.

96 See <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>.

97 See the OECD revised definition of "AI system": an "AI system" is "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment".

98 An "AI system lifecycle" involves the: "i) 'design, data and models'; which is a context dependent sequence encompassing planning and design, data collection and processing, as well as model building; ii) 'verification and validation'; iii) 'deployment'; and iv) 'operation and monitoring'. These phases often take place in an iterative manner and are not necessarily sequential. The decision to retire an AI system from operation may occur at any point during the operation and monitoring phase" – see OECD AI Principles (2019), section 1.1.

99 See <https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence>.

100 So far, some domestic AI regulations, such as the EU's AIA and Brazil's draft Senate Bill n. 2338/2023, seem to have adopted, almost verbatim, the OECD Principles definitions, including that of "AI system".

101 See ISO/IEC 22989:2022 (available at <https://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>).

102 See <https://wp.oecd.ai/app/uploads/2021/06/G20-AI-Principles.pdf>.

103 It states in this respect that it "does not have the ambition to provide one single definition of AI, since such a definition would need to change over time, in accordance with technological developments. Rather, its ambition is to address those features of AI systems that are of central ethical relevance". Yet, the Recommendation does provide a broad understanding of what "AI systems" mean, i.e., "systems which have the capacity to process data and information in a way that resembles intelligent behaviour, and typically includes aspects of reasoning, learning, perception, prediction, planning or control" (paragraph 2). For UNESCO, such broad understanding is "crucial as the rapid pace of technological change would quickly render any fixed, narrow definition outdated, and make future-proof policies infeasible" (UNESCO, 2023).

104 See <https://g20.utoronto.ca/2020/2020-g20-digital-0722.html>.

105 See http://www.g20italy.org/wp-content/uploads/2021/11/Annex1_DECLARATION-OF-G20-DIGITAL-MINISTERS-2021_FINAL.pdf.

106 See https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/390de76d-d4f5-4f45-a7b4-f6879c30c389/0fbffe8a/20231201_en_news_g7_result_00.pdf.

107 See <https://www.gov.uk/government/publications/ai-safety-summit-2023-chairs-statement-state-of-the-science-2-november/state-of-the-science-report-to-understand-capabilities-and-risks-of-frontier-ai-statement-by-the-chair-2-november-2023>.

108 See https://assets.publishing.service.gov.uk/media/66f5311f080bdf716392e922/international_scientific_report_on_the_safety_of_advanced_ai_interim_report.pdf.

109 See <https://aiforgood.itu.int/>.

110 See <https://unesdoc.unesco.org/ark:/48223/pf0000375322/PDF/375322eng.pdf.multi>.

111 See <https://aiforgood.itu.int/about-ai-for-good/un-ai-actions/unido/>.

112 See <https://aiforgood.itu.int/about-ai-for-good/un-ai-actions/wbg/>.

113 See <https://aiforgood.itu.int/about-ai-for-good/un-ai-actions/>.

114 See https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf.

115 See <https://www.un.org/techenvoy/global-digital-compact>.