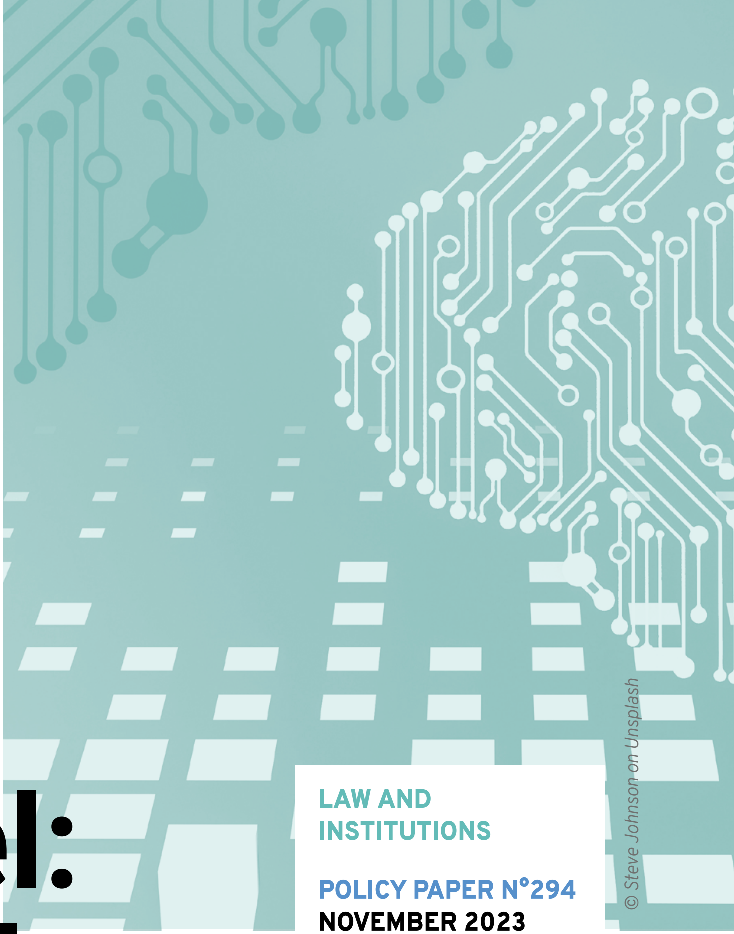
NOTRE EUROPE
Jacques Delors Institute
Penser l'Europe • Thinking Europe • Europa denken



© Steve Johnson on Unsplash

Regulating artificial intelligence at the EU level: obstacles and prospects

LAW AND INSTITUTIONS

POLICY PAPER N°294
NOVEMBER 2023

#artificialintelligence
#obstacles
#prospects
#AIregulation

Executive summary:

This paper aims to examine the obstacles and prospects for regulating artificial intelligence at the EU level. Firstly, a number of obstacles remain, including the difficulty of defining AI and the appropriate regulatory scope, the continued sway exercised by lobbying groups, along with the velocity of change in the AI industry which makes it challenging for regulators to keep-up. Secondly, in terms of the AI Act itself, the EU has chosen an approach that exhibits many strengths, relying on a 'technology-neutral' definition and setting out a 'risk-based' approach whereby AI systems are regulated according to the degree of risk they pose to society (the four categories include 'unacceptable risk', 'high-risk', 'limited-risk' and 'low or minimal risk'). The legislation still suffers from a number of inadequacies, however. These include insufficient flexibility in adapting to the speed of evolution in this sector, an over-emphasis on individual risks and thus weaker consideration of the broader societal-level impacts, and inadequate compliance frameworks which often rely on self-assessment. Thirdly, the EU is well positioned as a 'first mover' in the field of AI regulation to play a key role in influencing both national rules and international standards. Due to the complex and multi-faceted nature of AI technologies, the EU should consider a model such as the international regime for the prohibition of chemical weapons, with an additional forecasting unit, to establish global rules and monitoring of AI.

Arnault Barichella
is an affiliated Researcher at the Jacques Delors Institute, specializing in artificial intelligence, digital technologies and cybersecurity. He is also a post-doctoral researcher at the Université Paris-Saclay, and received his PhD in Political Science from Sciences Po Paris, where he remains affiliated with Sciences Po's Center for European Studies. Arnault obtained his Masters' degree from Sciences Po, his BA degree from Oxford University, and was awarded a visiting fellowship at Harvard University.

Introduction

Although technologies relating to artificial intelligence (AI) have been developing for several decades, the last year has seen an exponential acceleration which has propelled the topic front and center across the world. This is mostly due to the release of *ChatGPT* (version 3.5) by OpenAI, a California-based technology company, in November 2022. *ChatGPT* responds to brief prompts by generating text at a high speed, pulling together data collected from the Internet. While the accuracy and quality of AI-generated text has been the subject of much debate and criticism, this new technology has certainly succeeded in captivating the world's attention. *ChatGPT* has broken all records by becoming the fastest growing software and consumer application in history, winning over 100 million users globally in the space of only a few months.[1] Such dizzying success has triggered a global race with other tech companies, which have rushed to release their own AI chatbot models to compete with OpenAI, models ranging from Microsoft's *Bing* to Google's *Bard*.

This has contributed to launching what may be described as a 'civilizational change', ushering in a new industrial revolution or 'AI Age'. Some have welcomed these developments, praising AI's abilities to boost economic growth by enhancing prediction, improving resource allocation and personalizing services. AI can play a beneficial role in key sectors like healthcare, potentially helping to find remedies to diseases like cancer. At the same time, however, other voices have expressed alarm over the potential dangers associated with these new technologies. At one end of the spectrum, some scientists have warned of an existential threat to humanity, with the possibility of AI eventually becoming conscious, turning against its creators and choosing to eliminate the human species, much like what has been depicted in science fiction movies. More moderate and reasonable criticism has focused on the potential impacts of these technologies on employment. For example, Goldman Sachs has predicted that up to 300 million jobs could disappear internationally; the new jobs created in the AI industry may not compensate fast enough for those lost to automation over the next few years due to rapid advances in this field.[2] Other criticism has concentrated on the risk that chatbots present to democratic processes, with the possibility of amplifying online disinformation and manipulation through AI-generated texts, including its misuse by countries like Russia against the West, for instance.

More generally, concerns have been raised about AI software jeopardizing certain fundamental rights such as freedom of expression, non-discrimination and human dignity, as well as privacy and data protection.[3] In response to these rising apprehensions, many senior figures in sectors ranging from politics to business to civil society (along with people from the tech industry itself, including *ChatGPT* founder Sam Altman), have emphasized the urgent need for governments to regulate the AI industry at both the national and international levels. Last March, entrepreneur Elon Musk circulated an open letter, co-signed by other senior figures such as Apple co-founder Steve Wozniak, calling for a pause in the development of AI, arguing that the latter was not in the interest of humanity and risked rendering humans 'obsolete' in the near future.

[1] Hu K. (2023, February 2), *ChatGPT sets record for fastest-growing user base*, Reuters: https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

[2] At the same time, however, Goldman Sachs underlines that global productivity could increase by up to 7% during the same period due to advances in generative AI. See: Hatzius *et al.* (2023).

[3] In this regard, Italian regulators imposed a temporary ban on *ChatGPT* last April due to concerns over the software's unlawful collection of users' personal data.

Faced with this situation, the EU has responded rigorously, with the European Parliament approving last June the world's first ever comprehensive legislation to regulate artificial intelligence, adopting a risk-based approach that seeks to address many of the aforementioned concerns. While the EU's Artificial Intelligence Act (AIA) is not perfect, it still represents a genuine attempt by public authorities and democratic processes to take back control over an industry some fear is rapidly spinning out of hand. The AIA has the potential to become a global benchmark that could influence AI rules in many other countries around the world. This paper aims to analyze the obstacles and prospects for regulating artificial intelligence at the EU level, and how Europe has the potential to become a global norms setter in this area, as it already is in fields like climate change.

I . Obstacles to regulating AI in the EU and in general

There are a number of significant obstacles to the successful and effective regulation of AI at the EU level which have hampered progress over the last few years. First, as in other fields, strong lobbying pressure exercised by large corporations, especially technology firms forming part of the so-called GAFAM (Google, Amazon, Facebook, Apple and Microsoft), has rendered it more difficult to enact sufficiently robust regulation. Tech lobbying threats include 'pulling out from Europe', which the head of OpenAI Sam Altman said might happen since it would prove challenging for *ChatGPT* to comply with many aspects of the EU's AI Act. This could affect Europe's economic competitiveness vis-à-vis other major powers like the US and China, highlighting the need for strong international collaboration on the establishment of global AI standards and rules. What is more, the nomination by the European Commission of US economist Fiona Scott Morton to a key position, as one of the Chief Economists of the Directorate-General for Competition, triggered criticism due to her ambiguous ties with the American tech companies she would have been responsible for regulating. While Fiona Scott Morton eventually chose to withdraw her candidacy, this still raised concerns from Member States about the extent of lobbying pressure and the influence of the tech industry within EU institutions themselves.

Second, another seemingly trivial, but in reality, very significant obstacle has been the difficulty in agreeing upon a common definition of what actually constitutes an AI technology. This is important, because the nature of the definition of AI is essential in order to establish the ambit and format that regulation will subsequently adopt. Small differences in definition, including the fact that various technical terms may overlap with one another, will then carry major ramifications for how regulation can be developed and enacted. For instance, some technologies that are commercially marked as AI are quite simplistic and could easily be designated as statistical tools, whilst an overly narrow definition risks missing important new developments in generative AI, such as *ChatGPT*. Likewise, AI-enhanced video games should be regulated very differently from AI used for critical infrastructure. As a result, since AI can take on such diverse forms and possesses multi-pronged capacities, a 'one-size fits all' approach is problematic; a singular approach runs the risk of over-regulating certain technologies, whilst under-regulating other sectors. Thus, the EU has had to engage in a delicate balancing act to choose a specific definition and regulatory approach for the AI Act that is 'risk-based' and targeted (analyzed in more detail below).

Third, perhaps the greatest challenge of all has been the speed at which AI technologies continue to progress, making it very difficult for regulation to keep up with new developments happening on an almost daily basis. Until the release of *ChatGPT* in November 2022, AI had mostly been a tool used by software engineers;

ChatGPT changed this by turning AI into a consumer-focused software that ordinary people can utilize without requiring any technical expertise. It took OpenAI only four months to launch an upgraded fourth version of their product (*ChatGPT version 4*), with enhanced performances and a higher ratio of parameters. The latter became the fastest growing software in history, with more than 100 million users in the space of only two months, triggering a race amongst tech companies to develop their own large language models (LLM). Shortly thereafter, Microsoft chose to invest $13 billion in OpenAI, and incorporated the technology into many of its products, including a competing LLM called *Bing*. Google followed suit in early 2023 with the launching of its own LLM dubbed *Bard*, which sought to undercut competition by upgrading the AI software to produce texts on current events, which *ChatGPT* was initially unable to do since its data set extended only up to 2021 (this has been changed in subsequent upgrades to *ChatGPT* in Fall 2023). The CEO of Meta, Mark Zuckerberg, also joined the race, announcing that AI will become the single largest investment for his company and that it would be incorporated into every one of their products.

These examples highlight the acceleration of AI technological developments over the last year, which poses a fundamental challenge for regulators to keep-up. EU legislative processes move far more slowly. While the Commission submitted a proposal for the AI Act in April 2021, it was only ratified by the EU Parliament in June 2023, with final agreement expected in late 2023 or early 2024, followed by an implementation period of around eighteen months before all of its rules become fully binding and enforceable. Clearly, the slow pace of democratic consultation makes it very difficult to keep pace with the velocity of change in the AI industry. Due to the globalized nature of digital technologies, international regulation will also be necessary. The EU has been by far the fastest large political entity to enact any regulation on AI at all, with the US and China lagging behind.[4] This is compounded by the fact that existing regulatory systems for countries around the world were developed based on industrial era expectations about 'command and control'. Such expectations were already overtaken by the speed of technological developments for digital platforms over the last two decades. As a result, current legal frameworks are not sufficiently agile or responsive to address technological evolution in the AI era. While the first wave of the industrial revolution focused on enhancing human physical power, the AI revolution will continue to enhance human cognitive powers. Therefore, only a new regulatory approach will make it possible to keep pace; the extent to which the EU's AI Act succeeds in establishing such a new approach will be analyzed in the sections below.

II • Strengths and summary of the EU's AI Act

Policy-makers in the EU have sought to develop their concept of a 'human-centric' approach to AI, in order to ensure that European citizens will both benefit from these new technologies and be protected from the potential risks, ensuring that AI operates according to the EU's principles and values. This was highlighted in the 2019 Ethics Guidelines for Trustworthy AI and Policy, followed by the 2020 White Paper on Artificial Intelligence. Although EU institutions had initially set out a 'soft law' or non-binding approach through these two policy papers, recent developments made it clear that this would be insufficient. This led the EU to transition towards a legislative approach, with the aim of enacting a series of harmonized rules for the

[4] It should be noted that both China and the US have acknowledged the need to enact legislation on AI. Beijing has even expressed interest in the EU's AI Act, and may seek to transpose certain elements within its own domestic single market. In the US, while extensive discussions have been ongoing over the last couple of months, it is likely to take some time before Congress is able to reach consensus on any type of legislation for AI, which will in all likelihood be less ambitious than its EU counterpart.

development, market placing and utilization of AI software. The AI Act, approved by the European Parliament in June 2023 in its initial form, represents the culmination of such efforts.

The EU has chosen the format of a Regulation over that of a Directive, a significant choice since the law will be immediately binding on all Member States after its entry into force (instead of the longer process of transposing a Directive into national law). This arguably constitutes one of the strengths of the EU's approach, since the pace of change in the AI industry necessitates regulators to take rapid and decisive action. The objective of the EU's Artificial Intelligence Act is to ensure the continued effective operation of the Single Market by generating optimal conditions for the development and utilization of trustworthy AI technologies, harnessing AI's potential benefits whilst protecting society from associated risks. The Act seeks to establish a harmonized legal paradigm for the creation, placing onto the market and subsequent utilization by customers of AI services and products, including the ways in which ex-post controls will be conducted. At the heart of the Act is a 'technology-neutral' definition of AI systems, buttressed by a so-called 'risk-based approach'.

As previously explained, the definition of what in fact constitutes artificial intelligence remains highly contentious, with no single universal definition yet to be accepted by the scientific community. In response, the EU has opted for a compromise solution involving a 'technology-neutral' definition that has the merit of bringing clarity and simplification, when compared to more technical definitions. Drawing inspiration from the OCED's definition, the EU Act thus defines AI in Article 3(1) as: "… software that is developed with [specific] techniques and approaches [listed in Annex 1] and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

Annex 1 of the AI Act proposes a list of techniques currently used to develop AI, which encompass "logic and knowledge-based" systems, "machine learning", as well as various "statistical" approaches. One of the main advantages of this definition is arguably that it has a broad ambit, which can be used on a stand-alone basis, or as a component of a service or product. Another asset is that several provisions have been introduced to make the law 'future-proof' by seeking to encompass current and future AI technological developments. Hence, the adoption of delegated acts (Article 4) will become the primary tool for complementing the Annex 1 list with novel techniques and approaches relied upon to develop AI software as they emerge in the coming years. Moreover, the general AI definition is supplemented by an extensive list of other, more specific and technical definitions in Article 3. This includes defining a "provider" and "user" of AI technologies covering both public and private entities, along with definitions for a "distributor" and "importer" of AI, "biometric categorization" or "emotional recognition", amongst other definitions.

Perhaps the greatest strength of the EU's AI Act is the adoption of a so-called 'risk-based' approach, a choice criticized by some, but generally supported by politicians, academics, civil society and the business community. A risk-based approach involves legal intervention which is specifically adapted and tailored to the level of risk posed by different types of AI technologies. The Act thereby distinguishes between four different levels of risk; AI systems will be regulated only as strictly as is needed to address their specific risk level (from the highest to the lowest): (i) unacceptable risk, (ii) high risk, (iii) limited risk, and (iv) low or minimal risk.

1. **Unacceptable risk leading to prohibited AI practices:** This section is governed under Title II (Article 5) of the Law, which explicitly bans a number of AI systems or practices within the EU Single Market that are identified as posing an unaccep-

table risk to the livelihood, safety and rights of European citizens. These include AI technologies which rely on harmful manipulative “subliminal techniques”, those which take advantage of particular vulnerable groups (with mental or physical disability for instance), those which engage in social scoring practices (as in China), and finally those which rely on “real-time” remote biometric identification processes (such as facial recognition) in public spaces under the context of law enforcement.

2. **High risk leading to tightly regulated AI systems:** This section is governed under Title III (Article 6) of the Law. So-called “high-risk” AI systems, which impact in a potentially significant way people’s safety and/or fundamental rights, are not banned completely, but subject to very strict operating rules. The Act makes a distinction between two different types of high-risk AI systems:

   A. Those falling under the scope of EU health and safety harmonization laws, along with systems relied upon as a safety component for a product or service.

   B. Systems deployed under eight specifically defined areas within Annex III, which the Commission will be able to update according to technological evolutions through delegated acts (covered under Article 7). This list includes:
   - Management of critical infrastructure
   - Vocational training and education
   - Law enforcement
   - Biometric identification, including the categorization of persons
   - Employment and working management
   - Administration of justice, together with democratic processes
   - Migration and asylum, along with border control
   - Access to and use of key public and private services or benefits

The AI Act establishes strict regulations for such high risk systems. Chief among them is a requirement for “ex-ante conformity assessments”, whereby an EU database managed by the Commission would register such systems before their placement on the Single Market. While a number of AI products and services will fall under the ambit of existing safety laws and third-party conformity mechanisms (such as healthcare services), many others are not covered by existing rules and will thus be required to organize their own conformity assessments (self-assessment). At the time of writing, only high-risk AI systems relied on for biometric identification will need to undergo a conformity assessment from a “notified body”.

The list of strict requirements these high-risk AI systems will have to comply with is quite broad, and includes technical robustness, risk management, data governance and training, human oversight, transparency, as well as cybersecurity (see Articles 8 to 15). These rules apply in particular to providers, importers and distributors of high-risk AI systems. For instance, providers located outside the EU will need to have an authorized representative within the EU in order to adhere to the conformity assessment and to set up a post-market monitoring framework, amongst other obligations.

3. **Limited risk leading mostly to strict transparency rules:** Such “limited risk” AI systems include technologies that interact directly with humans, such as chatbots (*ChatGPT*, *Bard*, *Bing*, etc.), as well as biometric categorization, emotion recognition systems, together with software that generates or manipulates audio, video and image contents (e.g., deepfakes). Instead of being prohibited or subjected to a variety of strict regulations as with the above categories, limited risk AI systems are subjected instead to specific transparency rules. These focus mostly on the obligation for providers, importers and distributors of these technologies

to develop mechanisms so that it becomes immediately possible to identity with high accuracy when a text, image or audio has been generated by an AI system, as opposed to a human. This will include strict "water-marking" rules, in addition to the registration of AI algorithms within an EU-wide database to ascertain in cases of doubt the exact origin of a text, image or audio.

4. **Low or minimal risk, with no mandatory obligations:** All AI systems not falling under the ambit of the three previous categories are considered "low or minimal risk", meaning they can be deployed in the EU Single Market without conforming to any additional legal requirements.[5] However, the AI Act has still developed so-called "codes of conduct" to encourage voluntary compliance with the rules established for high risk and limited risk AI systems, examined above.

The robustness of the EU's AI Act is also apparent with respect to governance and enforcement processes. Under the Act, each Member State must identify a current institution or create a new entity to serve as a "national supervisory authority" charged with monitoring enactment and application of the Law. Likewise, the AI Act will establish a "European Artificial Intelligence Board" made-up of representatives from the Commission as well as Member States, charged with monitoring implementation at EU level. In addition, market surveillance institutions at the national echelon will play a key role for monitoring compliance pertaining to high-risk AI systems. In order to execute their tasks, these different entities will be provided with access to AI source codes, and will be charged with enacting any corrective measures to restrict, withdraw or prohibit AI systems that do not conform with the Law, if they are seen to represent a risk to public health, safety or fundamental rights.

Another notable aspect and arguably one of the strengths of the Law is that it establishes significant fines for non-compliance of up to €40 million or 7 % of the total worldwide annual turnover for private companies, depending on the degree of infringement. This is significantly more than the current 4% penalty ratio under the GDPR framework. It highlights the EU's intention to send a signal about its resolve to establish a strong regulatory framework for AI, complemented by strict penalty requirements at the national level. Aware of criticism (see below) that overly burdensome sanctions might stifle innovation, the AI Act also proposes several corrective measures, including the establishment of a "regulatory sandbox". This involves the creation of a so-called "controlled environment" to enable the safe development and testing of new AI technologies for a designated period of time, before their commercialization. In addition, a number of measures are proposed to support innovation for small start-ups and providers, highlighting the EU's attempt to establish a balanced approach within the legislation.

III • Weaknesses of the EU's AI Act and policy recommendations

In spite of the many positive elements examined above, a number of weaknesses are apparent. Although certainly constituting the most ambitious legislative attempt to regulate AI in the world, the EU's legislation still has room for improvement. Firstly, despite several provisions designed to address this, it is unclear whether or not the AI Act will provide adequate flexibility in adapting to high-speed evolutions in this sector. In this regard, the current framework for adding new AI systems to the high-risk category listed in Annex III might not be sufficiently agile and open-ended. Hence, it could be useful to authorize the European Artificial Intelligence Board to propose changes to the Annex on a regular basis, including broadening the exis-

[5] For instance, most video games fall into this 'low or minimal risk' category.

ting categories over time; this might involve institutionalizing particular timeframes under which this would take place. Another improvement would be to enhance the flow of information between national governments and EU institutions by systematically examining and compiling incident reports from Member States.

Secondly, while the risk-based approach adopted by the Law has many benefits, it also results in an emphasis on individual risk, without sufficient consideration of the broader, societal-level impacts of AI. Like many other technological innovations, AI can result in societal harm, even when its direct impact on individuals remains minimal. For instance, AI may be used to create false online content or disinformation which reduces people's general trust in science; this is obviously detrimental to society, without causing direct harm to individuals. As a result, it would be beneficial to move beyond the current focus on AI risks to individual human safety, health and fundamental rights to include an evaluation of broader harms to society as a whole. This would involve connecting the AI Act to the EU's broader governance ecosystem to ensure that societal risks are fully taken into account. Different options might encompass a mandatory impact assessment with the possibility of a channel for societal feedback; better public monitoring frameworks to enhance the diffusion of AI's societal impact; along with the potential initiation of certain procedural rights with a societal dimension (access to justice, information, etc.), together with better public inclusion in AI governance at the EU level.

Thirdly, even on focus issues such as the protection of fundamental rights for individuals, the law's provisions may be insufficient to achieve the stated objectives. For instance, the definitions used for biometric categorization and emotion recognition might not be precise or inclusive enough, which could jeopardize fundamental rights protection. Likewise, it would be useful to enlarge the criteria for unacceptable risk to enhance agility in coping with rapid developments in the AI sector to include new prohibited practices which pose a threat to fundamental rights. Similarly, the AI Act focuses mostly on strict obligations for providers and perhaps not enough on the users of AI systems; thus, a common mandatory framework for impact assessment on all high-risk AI systems would help to correct this.

Fourthly, perhaps one of the greatest flaws in the EU's AI Act has to do with its compliance and enforcement systems. In its current version, the Law privileges self-assessment by private companies. As examined above, there are a number of exceptions to this, where more rigorous third-party assessments are carried out. Yet, this still raises questions about the legal enforceability of the Act in its entirety, since private companies will obviously have an incentive to self-report in such a way as to exaggerate their level of compliance with EU rules. Therefore, there is a clear need to restrict the categories of AI systems relying on self-assessment, and to enlarge those which must undergo third-party evaluation so that the latter becomes the norm over time.

Finally, over 150 CEOs and executives from large companies such as Renault, Heineken, Airbus and Siemens signed an open letter to EU institutions warning about their perceived concerns over the costs of compliance with the AI Act. They have highlighted that the Act risks making European economies uncompetitive, especially since actors like the US and China will have to comply with far fewer and less stringent regulations on AI when compared to Europe (see below). They point to the Center for Data Innovation (CDI) report which argued that the EU's AI Act will have a very negative impact on Europe's economy and competitiveness.[6] The report warned that the legislation could end-up costing the Union up to €31 billion over

[6] Mueller (2021).

the next five years, whilst reducing AI-related investments in the EU by nearly 20%. However, many other academic studies published since then have cast substantial doubt on this rather pessimistic analysis, arguing that the overall impact will in all likelihood be mixed, boosting competitiveness in some areas but not in others.[7] This is because the CDI's report did not take into consideration the indirect economic benefits that will accrue from the regulation to the public,[8] and relied on an incomplete methodology which has since then been largely refuted.

Nevertheless, many small and medium-sized enterprises (SMEs) in Europe are concerned they will be disadvantaged when compared to larger companies, in terms of dealing with compliance costs associated with the EU's AI Act. One way to (partly) address this issue might be to create new structures for associating SMEs more closely in the elaboration of standards for AI regulation. This could include adding an 'AI branch' to existing public-private partnerships on digital/cybersecurity issues between the EU and the private sector, such as the 'Contractual Public Private Partnership' (cPPP).[9]

IV • How the EU's AI Act can play a key role in setting global standards

Because the EU's AI Act will constitute the world's first comprehensive legislation to regulate artificial intelligence, it is already being presented as a potential benchmark for other countries to follow, and will aid in the establishment of international standards in this area. This can be termed as a 'first mover' advantage, whereby the first large entity to establish comprehensive regulations in a given field plays an essential role in shaping the rules and debate from that point onwards for all other nations. This is largely due to the interconnected nature of the globalized economy and of digital technologies, in particular. In this regard, it is worth mentioning that a parallel has been drawn with the passage of the EU's General Data Protection Regulation (GDPR) in 2016. At that time, the GDPR constituted the first concerted attempt to establish far-reaching rules to protect citizens' data and privacy; the GDPR subsequently became a global standard that inspired the adoption of similar rules by nations across the world. The same situation could potentially emerge from the EU's AI Act, since a number of Parliaments in other countries have already begun studying the initial version passed by the European Parliament as a (partial) blueprint. This includes democracies like Canada, Japan, Brazil, and the US to a certain extent, even though American fears of over-regulation and a culture of laissez-faire economics mean that only certain aspects of the EU's law are likely to be emulated in Washington. The same also applies to some non-democratic regimes like China, where Beijing has expressed interest in several aspects of the EU's AI Act, even though it is likely to be implemented very differently under the Chinese political system.

In other fields such as environmental protection, climate change and data protection, the EU has been able to rely on its impressive market power as the world's largest trading bloc to influence and exercise pressure over potential trading partners. Countries wishing to gain access to the EU's single market, or who aim to negotiate a bilateral trading agreement, have to comply with EU standards in these key areas. This has enabled Europe to have a positive influence in raising environmental or pri-

[7] Haataja and Bryson (2021), Center for European Policy Studies (2021).

[8] Heikkilä (2021).

[9] The 'Contractual Public Private Partnership' (cPPP) was established in 2016 and includes the European Commission together with the European Cyber Security Organization, bringing together public and private entities that work in partnership to reinforce cybersecurity for critical infrastructure.

vacy protection standards in many countries around the world. Europe remains the largest single market in the world, where consumers enjoy a high standard of living. For these reasons, the EU is arguably well positioned to exercise decisive influence over other major economic powers concerning the establishment of strict and rigorous AI standards. As with other issues, countries wishing to trade and gain access to the EU's single market will have to adopt many rules contained in the EU's AI Act. In fact, several clauses explicitly point towards its extraterritorial ambit. Once again, this resembles the GDPR, whose promulgation caused controversy at the time since the strict privacy protection rules it established apply to any entity, regardless of whether or not it is geographically located in Europe, as long as it engages in the collection of EU citizens' data. Likewise, the AI Act as passed by the European Parliament in June 2023 explicitly asserts that the new rules will apply not only to providers and users of AI software based within an EU Member State, but also to those located in a third country that are placing AI services or products onto the Single Market for usage within the EU.[10]

Moreover, in addition to influencing AI norms and standards in other countries, the EU has asserted its intention to play a key role in shaping the development of international regulations under the aegis of the United Nations or the OECD. Many commentators, including ChatGPT founder Sam Altman, have called for the establishment of a new international organization specifically dedicated to regulating AI; Altman points to the International Atomic Energy Agency's regulation of the peaceful use of nuclear energy as a model. With some scientists warning about the existential risk posed by AI technologies, a parallel between nuclear weapons and AI has often been drawn, along with the need to establish strict international rules to ensure AI's peaceful use, as was done with nuclear power at the end of WWII. There are certainly many relevant points of comparison to be drawn, and international AI regulation can find useful precedents in the global regime established for nuclear non-proliferation and the regulation of nuclear energy.

While some have pointed to international conventions regulating the use of dual use technology (i.e., technologies that serve a dual military and civilian function), these conventions make no explicit mention of machine learning software or AI. Therefore, while the latter could possibly be upgraded, the international treaty framework and Organization for the Prohibition of Chemical Weapons (OPCW) might be a more helpful point of reference for the emerging global AI regulatory regime. The fact that nuclear can only be used as a weapon or as a source of energy creates a rather stark dichotomy which lacks nuance, and may not be appropriate for a complex technology such as AI. Chemicals, however, can be used not only as weapons, but for a broad variety of uses and they are ubiquitous across society in industry, agriculture, healthcare, manufacturing or construction. Therefore, chemical regulation could serve as a better point of reference, since a number of analysts have pointed to the likelihood that AI will also become ubiquitous across society, impacting most sectors in a way comparable to chemicals. For these reasons, the global regime for the prohibition of chemical weapons, which adopts a more wholistic approach to the international regulation of chemicals, arguably provides a more appropriate framework to draw inspiration from. Nonetheless, any international agency tasked with monitoring AI will need to include a robust forecasting unit, staffed with international experts capable of anticipating rapid algorithmic developments and new AI applications. The EU should pursue a leading role in the development of such a global monitoring and forecasting body for the regulation of AI.

[10] There are certain exceptions to this however, since the current version of the EU's AI Act would not apply to AI systems created or utilized exclusively in the context of military operations, to public entities located in third countries, nor to international organizations or for purposes relating to international law enforcement and judicial collaboration.

Conclusion

In conclusion, this paper has examined different facets regarding the obstacles and prospects for regulating artificial intelligence at the EU level. A number of obstacles remain to regulating AI in the EU and in general, including the continued sway exercised by lobbying groups, the difficulty of defining AI and the appropriate regulatory ambit, along with the pace of change in the AI industry which makes it challenging for regulators to keep-up. In terms of the AI Act itself, the EU has chosen an approach that exhibits many strengths, relying on a 'technology-neutral' definition and setting out a 'risk-based' approach whereby AI systems are regulated according to the degree of risk they pose to society. Yet, despite these robust elements, the EU's AI Act suffers from a number of inadequacies, including insufficient flexibility to adapt to high-speed evolution in this sector, an over-emphasis on individual risks and thus insufficient consideration of broader societal-level impacts, along with inadequate compliance frameworks in some cases relying too much on self-assessment.

As a result, this paper has proposed a number of policy recommendations to address these weaknesses. This includes enhancing the flow of information between national governments and EU institutions by systematically compiling incident reports, as well as moving beyond the current focus on AI risks to individual human safety, health and fundamental rights to include an evaluation of broader harms to society as a whole. Likewise, it would be useful to enlarge the criteria for unacceptable risk to enhance agility in coping with rapid developments by including new categories of prohibited practices, which could take the form of a common mandatory framework for impact assessment on all high-risk AI systems. There is also a need to restrict the categories of AI systems relying on self-assessment and to enlarge those which must undergo third-party evaluation, so that the latter becomes the norm over time. Moreover, the creation of new structures for associating SMEs more closely in the elaboration of AI standards would help to mitigate adverse economic impacts from regulations. Finally, the EU is well positioned to mobilize its position as a 'first mover' in the field of AI to play a key role in influencing national rules in countries around the world, along with international standards in this field. Due to the complex and multi-faceted nature of AI technologies, it might be beneficial for the EU to rely on the international regime for the prohibition of chemical weapons as a model for establishing global rules on AI, instead of the global nuclear regime which tends to be overly narrow.

Regulating artificial intelligence is likely to become one of the most significant challenges facing the international community over the next few decades. AI will revolutionize most aspects of society and is likely to become ubiquitous across many sectors. It has the potential to bring real benefits to mankind, boosting economic development and possibly leading to major breakthroughs in key sectors like healthcare. At the same time, however, AI brings with it a host of major risks ranging from the threat of substantial job losses, an amplification of online disinformation or manipulation of democratic processes, along with threats to fundamental rights and potentially worse depending on how AI develops in the future. The EU's AI Act represents an ambitious attempt to establish a comprehensive regulatory regime, the first in the world with potentially major international ramifications. While this represents a positive first step, much more needs to be done in the near future to ensure that Europe and the world can harness the potential benefits of the AI revolution, while minimizing its associated risks.

Bibliography

- AccessNow (2021), *Access Now's submission to the European Commission's adoption consultation on the AI Act.*
- Biommasani R. *et al.* (2023), *Do Foundation Model Providers Comply with the Draft EU AI Act?*, Stanford University – Center for Research on Foundation Models.
- Bolkar S. (2023), *EU AI Act: The Regulatory Framework on the Usage of Machine Learning in the European Union.*
- Casillo K. and Powell A. (2023), *Challenges in regulating the use of Arti1cial Intelligence*, ENSafrica.
- Center for European Policy Studies (2021), *Clarifying the costs for the EU's AI Act.*
- Clark S. et al. (2021), *Submission of Feedback to the European Commission's Proposal for a Regulation laying down harmonized rules on artificial intelligence*, University of Cambridge – Centre for the Study of Existential Risk.
- European Digital SME Alliance (2021), *DIGITAL SME reply to the AI Act consultation.*
- European Parliament (2023), *EU AI Act: first regulation on artificial intelligence.*
- European Parliament (2023), *Artificial Intelligence Act*, Briefing EU Legislation in Progress.
- European Parliament (2023), *EU Digital Markets Act and Digital Services Act explained.*
- Feingold S. (2023, June 30), *The European Union's Artificial Intelligence Act – explained*, World Economic Forum.
- Future of Life Institute (2021), *FLI Position Paper on the EU AI Act.*
- Gibson Dunn (2023), *European Parliament Adopts its Negotiating Position on the EU AI Act.*
- Haataja M. and Bryson J. (2021), *What costs should we expect from the EU's AI Act?*, SocArXiv.
- Hatzius J. *et al.* (2023), *The Potentially Large Effects of Artificial Intelligence on Economic Growth*, Goldman Sachs – Global Economics Analyst.
- Heikkilä M. (2021, September 8), *Decoded: Parliament is back — The price of AI — Doctors > AI*, POLITICO.
- Hu K. (2023, February 2), *ChatGPT sets record for fastest-growing user base*, Reuters.
- Lekatis G. (2023), *The EU Artificial Intelligence Act*, Cyber Risk GmbH.
- Maquindus O. (2023, June 23), *On the back of its AI Act, the EU invests in risky software to secure borders*, Le Monde.
- McFadden et al. (2021), *Harmonizing Artificial Intelligence: The Role of Standards in the EU AI Regulation*, Oxford Commission on AI & Good Governance.
- Mueller B. (2021), *How Much Will the Artificial Intelligence Act Cost Europe?*, Center for Data Innovation.
- O'Shaughnessy M. (2022), *One of the Biggest Problems in Regulating AI Is Agreeing on a Definition*, Carnegie Endowment for International Peace.
- Rignell M. (2023), *Investing in safe and trustworthy AI: a European imperative, a French opportunity*, Institut Montaigne – Policy Paper.
- Smuha N. *et al.* (2021), *How the EU can achieve legally trustworthy AI: A response to the European Commission's proposal for an Artificial Intelligence Act*, LEADS Lab – University of Birmingham.
- Smuha N. (2021), "Beyond the individual: governing AI's societal harm", *Internet Policy Review*, 10(3).
- The Future Society (2021), *Proposal for a Regulation – "Artificial Intelligence – ethical and legal arguments": Trust in Excellence & Excellence in Trust.*
- Veale M. and Borgesius F. Z. (2021), "Demystifying the Draft EU Artificial Intelligence Act: Analysing the good, the bad, and the unclear elements of the proposed approach", *Computer Law Review International*, 4(21).
- Wheeler T. (2023), *The three challenges of AI regulation*, Brookings Institution.
- Ziady H. (2023, June 15), *Europe is leading the race to regulate AI. Here's what you need to know*, CNN International.

Managing Editor: Sylvie Matelly • The document may be reproduced in part or in full on the dual condition that its meaning is not distorted and that the source is mentioned • The views expressed are those of the author(s) and do not necessarily reflect those of the publisher • The Jacques Delors Institute cannot be held responsible for the use which any third party may make of the document • Original version • Edited by Marjolaine Bergonnier • © Jacques Delors Institute

Institut Jacques Delors
Penser l'Europe • Thinking Europe • Europa Denken
18 rue de Londres 75009 Paris, France • www.delorsinstitute.eu
T +33 (0)1 44 58 97 97 • info@delorsinstitute.eu

NOTRE EUROPE
Institut Jacques Delors
Thinking Europe • Penser l'Europe • Europa denken

PREMIER MINISTRE
Liberté
Égalité
Fraternité

Ce projet reçoit des financements du programme Citizens, Equality, Rights and Values Programme (CERV) de la Commission européenne sous le numéro Project 101104850 – IJD 2023.