DODS | POLITICAL INTELLIGENCE

Understanding the EU's risk-based approach to regulating Artificial Intelligence

By Eleonora Bottin

Introduction

The rapid development of Artificial Intelligence (AI) has presented policymakers across the world with an urgent challenge to work out how to regulate the technology in a way that enables it to flourish but minimizes its risks to society. Against this backdrop, the European Union is blazing a trail with its planned AI Act, the first law on AI by a major regulator. After initially taking a hands-free approach to the technology, the European Commission introduced the landmark legislation that promises to classify AI applications and systems according to risk and regulate accordingly. The law, which the EU hopes to adopt by the end of 2023, would ban the use of AI in intrusive biometric surveillance, require greater transparency from generative systems like ChatGPT, but have minimal requirements on most AI applications such as email spam filters. The EU is also reaching out to coordinate with other major regulators, in particular the U.S., to ensure the global regulatory framework for the technology that transcends national borders reflects its democratic and human rights values.

This Dods EU Political Intelligence report provides a one-stop guide to the EU's approach to regulating AI, including an overview of the evolution of its policy, the key elements of the landmark AI Act, cooperation with the US, and a selection of stakeholder reactions.

Evolution of the EU approach to AI

The European Union's Artificial Intelligence Act, a key component of its wider digital agenda for Europe, promises to be the world's first law for the fast-evolving technology. The landmark regulation represents the culmination of a strategy to build a legal framework which enables AI to flourish in the EU, but also protects Europeans from the risks associated with certain uses of the technology.

At a European Council in October 2017, EU leaders agreed the bloc needed a sense of urgency to address emerging trends such as AI and blockchain technologies "ensuring a high level of data protection, digital rights and ethical standards", and invited the European Commission to put forward an approach to AI by early 2018. The Commission responded with a communication on Artificial Intelligence for Europe in April 2018, which recognised that AI was already in use in many forms, from phones suggesting songs to virtual personal assistants, and was helping solve major challenges from tackling chronic disease and fighting climate change. Amid fierce global competition, a solid European framework was needed, it said, adding that the EU approach would capitalise on its research strengths, digital single market, and wealth of data to feed AI systems.

To inform its approach, the Commission launched the European AI Alliance in June 2018 to engage with stakeholders, including civil society, business organisations, and academia, through a series of events, public consultations, and online forums. It also established a high-level expert group on AI. In December of the same year, the Commission published its Coordinated Plan on Artificial Intelligence, providing guidelines for its stated ambition for Europe to become the world-leading region for developing and deploying "cutting-edge, ethical and secure AI". It followed that up in June 2019, by adopting a communication on building trust in what it called human-centric artificial intelligence.

The European Commission initially leaned towards a light-touch approach, with its non-binding April 2019 Ethics Guidelines for Trustworthy AI and Policy and investment recommendations for trustworthy artificial intelligence. However, it soon looked to formulate an EU-wide legislative approach to the use of AI, a move also seen in other jurisdictions such as the US and UK. It

launched a white paper on AI in February 2020, followed by further consultation and impact assessment of its planned legislative approach, a process which culminated in the publication of its proposal for an Artificial Intelligence Act in April 2021, together with a revised coordinated plan. A couple of months later at their June 20 Council meeting, EU leaders acknowledged that AI offered great benefits but also risks, and called on the Commission to put forward concrete proposals "which follow a risk-based, proportionate and, if necessary, regulatory approach for artificial intelligence." The latest component of the Commission's strategy is a proposal for an AI Liability Directive presented in September 2022.

The European Parliament has also played a major role in developing the AI regulations. In a February 2017 resolution, it said it was vitally important for the legislature to consider the legal and ethical implications and effects of new sophisticated robots and other manifestations of AI, without stifling innovation. More recently, the creation of the Parliament's Special Committee on Artificial Intelligence in a Digital Age (AIDA) in June 2020 was another important step. The Committee had an initial one-year-mandate with the objective of studying the challenges and effects of AI as well as providing recommendations for determining the EU's objectives. In March 2022, AIDA adopted its report and at the beginning of May of the same year, the Parliament agreed on its final recommendations. MEPs called for an EU Roadmap to 2030 and underlined the potential benefits of the technology and the need for Europe to act as a "global standard-setter". In addition, the Parliament addressed the risks to fundamental rights and privacy, especially through mass surveillance.

Artificial Intelligence Act

The Commission's first-of-its-kind legislation aims to ensure that AI systems in the EU market are safe and respect the union's values, provide certainty to facilitate investment and innovation in the bloc, and enhance governance and enforcement of AI application. To achieve this, it promises to regulate AI systems and applications according to the risks they pose, dividing them into four categories: unacceptable risk, high-risk, limited risk, and minimal risk.

Unacceptable risk AI systems which are considered a clear threat to safety, livelihoods, and the rights of citizens will be banned. At the time of writing, these would include systems that involve so-called social scoring, the classification of people based on behaviour, personal characteristics, or economic status; cognitive behavioural manipulation of people or vulnerable groups; and real-time and remote biometric identification systems like facial recognition, except for delayed use by criminal justice authorities with court approval.

The high-risk category, which includes systems which could harm safety or fundamental rights, will be subject to strict controls before entering the market and throughout their lifecycle. They will face risk assessment, human oversight, traceability measures, and compliance with a high level of robustness and security. The Commission's original proposal for what should be included in the high-risk category has been the focus of intense debate in both the Council and European Parliament. The Commission identified as high risk the systems used in the following sectors: critical infrastructures, educational or vocational training, safety components of products, employment and management of workers, essential private and public services, law enforcement, migration, asylum, and border control management, and administration of justice and democratic processes.

By contrast, limited risk AI systems need only comply with transparency rules to ensure users are aware of the use of AI and can make informed decisions. These would apply to applications that

generate or manipulate images, audio or video content, such as deepfakes, or programs such as chatbots, so users realise they are interacting with a machine. Similarly, generative AI systems, like ChatGPT, would also have to disclose that the content is produced by AI and ensure that it was designed to not generate illegal content. The AI Act would not impose any restrictions on systems deemed to pose minimal risk, which includes most AI applications used in the EU today such as spam filters or AI-enabled computer games.

The Council of the EU reached is general approach to the new legislation on the 6 December 2022 under the Czech Presidency. The member states stressed the importance of having AI systems which are safe and respect existing rules on fundamental rights and Union values. In its general approach the Council narrowed down the definition of AI to those systems that are developed through machine learning and logic- and knowledge-based approaches. The Council was responsible for adding social scoring to the list of banned systems and making an exception for law enforcement to use biometric identification. The member states also clarified the requirements for high-risk systems to reduce the burden for providers and provided more detail on the allocation of responsibilities along value chains. The Council introduced new provisions concerning how the legislation would apply to general purpose AI, while national representatives worked on several other aspects of the text including provisions concerning the conformity assessment, the autonomy of the AI Board, penalties, and transparency. Lastly, the Council modified substantially the measures concerning the support of innovation, a key concern of member states.

The legislation has been the subject of intense negotiations among the political groups in the European Parliament, with the MEPs presenting more than 3,000 amendments to the draft text. The final political agreement on the text was reached on April 27, 2023 and then on 10 May, the Committee on Internal Market and Consumer Protection (IMCO) and Committee on Civil Liberties, Justice and Home Affairs (LIBE) adopted their report on the AI Act. The parliamentary rapporteurs, Brando Benifei (S&D, IT) and Dragos Tudorache (RE, RO), stressed the importance of the file for the EU and internationally, and said it was crucial to put fundamental rights at the heart of the legislation. In keeping with efforts to ensure the EU's approach chimes with other major regulators, MEPs agreed a definition of AI in line with the OECD. Most of the EU's member states signed up to the OECD's principles on AI in 2019 which aim to promote AI which is innovative and trustworthy and that respects human rights and democratic values. The MEPs also amended the original proposal to include bans on intrusive and discriminatory uses of AI as well as a full ban of AI instruments for biometric surveillance, emotion recognition, and predictive policing, following concerns from S&D, Greens/EFA, and Renew. EPP lawmakers pushed for amendments for two exemptions concerning security: missing children and terrorism, but these were rejected. On the 14 June, the European Parliament adopted the AI Act report with a large majority.

In a sign of the determination to get the landmark law over the line, the co-legislators and the Commission held the first trilogue meeting to discuss the legislation on the same day as the Parliament's plenary vote, with Benifei explaining lawmakers were pushing to complete the legislative process before the end of the current parliamentary term. The rapporteurs, reporting back to the LIBE and IMCO Committees, noted that the Commission and Council showed some openness to discuss the Parliament proposals, including impact assessments to ensure the responsible and ethical use of AI systems. The second trilogue took place on 18 July, with additional meetings scheduled for 26 September and potentially 26 October 2023. So far, the co-legislators discussed the following topics: notifying authorities and notified bodies (art. 30 to 39), obligations of providers and users of high-risk AI systems and other parties (art. 16 to 27), and

standards, conformity assessment, certificates, registration (art. 40 to 51). The lawmakers are aiming to conclude the negotiations by the end of the year.

Věra Jourová, the European Commission's Vice-President for Values and Transparency, during a European Parliament Plenary meeting on 13 July 2023, said that the Commission wanted voluntary commitments from companies to start implementing the AI Act ahead of its legal deadline, adding that they could not wait until the legislation became applicable in two years due to the rapid development of the technology. She added that the EU was engaging internationally at with several organisations including the OECD, United Nations, Council of Europe and bilaterally, including the G7's Hiroshima AI process to develop guardrails for generative AI. Jourová said the EU would also work with the UN on the global digital compact, and more specifically with UNESCO on the recommendations on ethics of AI. However, some MEPs voiced concerns about the voluntary agreements and the AI Pact, saying it should not undermine the work done on the AI Act or influence the trilogues.

Artificial Intelligence Liability Directive

On the 28 September 2022, the Commission presented a proposal for a new Artificial Intelligence Liability Directive, the first time it has presented rules for damages caused by AI systems, and also proposed a revision of the EU Product Liability Directive to modernise the existing rules on liability of manufacturers for defective products. The new rules will set out measures for any type of victim to have a fair chance of receiving compensation if they are harmed by a provider, developer, or user of AI. The AI Liability Directive alleviates the victims' burden of proof by introducing the 'presumption of causality.' If victims can prove the non-compliance of AI products with a certain obligation relevant to harm, a court can presume that this non-compliance caused the damage. The proposal also introduces the right of access to evidence. Victims will be able to ask a court to order the disclosure of information about high-risk AI systems, which will allow victims to identify the person that could be held liable. Consequently, victims will have right to compensation if the damage it is proven. At the time of writing, this proposal was in the early stages of the legislative process and expected to develop further in the months ahead.

EU-US Trade and Technology Council

The EU's approach to AI has also included a concerted effort to coordinate and shape the agenda with other key national regulators, in particular the US. AI has become a key focus of the EU-US Trade and Technology Council (TTC), a bilateral forum set up in mid-2021 which aims to help deepen transatlantic trade and economic relations based on shared values. The TTC published a Joint Roadmap on Evaluation and Measurement Tools for Trustworthy AI and Risk Management on 1 December 2022, underlining their commitment to cooperate on standards, tools, and methodologies and develop common terminology and taxonomy. However, the EU and the US will maintain complete autonomy over their respective policy and legislative measures. The two sides have voiced different views on certain issues, including the EU's classification of AI systems in different "risk" categories. During a stakeholders' event in February 2023, an official from the US National Institute of Standards and Technology (NIST), a leading voice in the development of AI standards, said that risk tolerance and threshold should be application-based and its management should not be reduced to a checklist.

At the fourth TTC ministerial meeting on 31 May 2023, the two sides said they had agreed to launch three expert groups to work on AI terminologies, taxonomies, standards and emerging risks, focus on generative AI systems, and prioritise parts of the Administrative Arrangement on

Artificial Intelligence for the Public Good signed in January to address global challenges on climate change, natural disasters, healthcare, energy and agriculture. The transatlantic partners also reaffirmed their commitment to a risk-based approach to AI. The White House said, the working groups issued a list with 65 AI terms, crucial for the understanding of the risk-based approach, along with both parties' interpretations of the definitions. The EU and the US also reaffirmed their commitment to cooperate and share research on AI.

Stakeholder reactions

The rapid spread of AI across a range of sectors has resulted in an equally varied response to the AI Act. In general, stakeholders have welcomed the need to regulate AI and the EU's values-based and human-centric approach. However, several have voiced concerns about the complexity of some of the elements introduced by the European Parliament, the risk of limiting innovation and technological development, and highlighted gaps in human rights protections.

Amnesty International and EDRi, a network of groups and experts working to defend digital rights, welcomed the full ban on abusive mass surveillance technology and on emotional recognition in key sectors like education, law enforcement, and the workplace. However, Amnesty International expressed concerns about a lack of measures where AI systems affect migrants, refugees, and asylum seekers. Consumer organisation BEUC said it regretted that the Parliament proposals would give businesses the option to decide if their AI system was considered high-risk or not, and thus escape from the resulting legal consequences.

Some organisations have raised questions about what systems will ultimately be included in the high-risk category. The Information Technology Industry Council (ITI) said it was concerned about a departure from the risk-based approach in the provisions on foundation models and the expanded list of high-risk AI uses. It added that the trilogue negotiations would be fundamental to address these issues. The Computer and Communications Industry Association (CCIA) also expressed concerns on the Parliament's decision to extend the strict requirements meant for high-risk cases to some other AI systems that it argued pose very limited risks. These stakeholders also stressed the importance of the AI Act focusing on fostering innovation. The Developers Alliance argued that the Parliament plans risked making it more complex and could discourage development, and called for a more "feasible" approach.

The growing use of generative AI systems has raised questions about copyrights issues. The Society of Audiovisual Authors (SAA) said that while the Parliament had taken a positive stance on some aspects of generative AI that were not in the Commission's original plans, its proposals do not adequately address the challenges faced by creators. There was uncertainty on the legal status of content generated by an AI which could have been ingested from other sources or mimicked creative work of others, it said. The SAA argued that the EU's Copyright Directive is not sufficient to address this new technology. ICMP, the global trade body for music companies, composers and songwriters, urged the EU lawmakers to be ambitious and conclude a "gold-standard law" which would enable AI to flourish, ensure more detailed reporting obligations, set suitable sanctions for breaches and recognise copyright as the fundamental catalyst for music creation, commerce and compensation. And European Visual Artists (EVA) group said transparency obligations on generative AI must ensure that works are used with the consent of the rightsholders and descriptive enough to provide information on who and how works are used.

Eleonora Bottin is a Dods EU political consultant covering digital policy.

Register here to join the free Dods Political Intelligence webinar with leading AI experts:

- Lord Clement-Jones, co-chair of the UK All-Party Parliamentary Group on AI
- Sana Khareghani, former Head of the UK Office for AI
- Adriano Koshiyama, co-founder of AI start-up Holistic AI
- Florian Ostmann, Head of AI Governance and Regulatory Intervention at The Alan Turing Institute

Key questions for the webinar include:

- What are the biggest AI risks and rewards that UK and other regulators must consider?
- Has the UK got the right balance between AI risk and reward with its pro-innovation approach?
- Does the approach to AI regulation in the UK compare favourably with that of the EU and other major regulators?
- What can we reasonably expect the UK prime minister's planned AI summit to deliver?