

Oldenburg, Derk; Serentschy, Georg

Conference Paper

How to regulate AI?

32nd European Conference of the International Telecommunications Society (ITS): "Realising the digital decade in the European Union – Easier said than done?", Madrid, Spain, 19th - 20th June 2023

Provided in Cooperation with:

International Telecommunications Society (ITS)

Suggested Citation: Oldenburg, Derk; Serentschy, Georg (2023) : How to regulate AI?, 32nd European Conference of the International Telecommunications Society (ITS): "Realising the digital decade in the European Union – Easier said than done?", Madrid, Spain, 19th - 20th June 2023, International Telecommunications Society (ITS), Calgary

This Version is available at:

<https://hdl.handle.net/10419/278007>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

ITS European Conference 2023

Paper Derk Oldenburg and Georg Serentschy¹

09.06.2023

How to regulate AI?

Introduction

Regulating emerging technologies has always been a challenge. Balancing the risks and opportunities of AI is raising the bar, in particular since generative AI systems such as ChatGPT have been made accessible to a large public. The rapid development of AI deserves another look at what balancing means and how we could look at it in the face of the challenges confronting our society.

The debate and the EU AI Act

The ‚Hiroshima AI Process‘, launches an effort of the G7 countries to ‘Advance international discussions on inclusive artificial intelligence (AI) governance and interoperability to achieve our common vision and goal of trustworthy AI, in line with our shared democratic values’². The fact that AI governance has moved up the ladder of G7 priorities this quickly illustrates the intensity of the debate about the impact of AI on the economy, society and on our daily lives.

The communique of the value-driven leading industrial countries also mentions a number of global challenges for which, for quite some time, hopes have existed that AI could be part of the solution. The text mentions, inter alia: economic resilience, decarbonization & climate crisis, food security, global health and migration.

These hopes rightly still exist and with the rapid development of the technology, the contours of the benefits and opportunities of AI have become sharper and their scope has increased. Existing applications with a specific purpose, known as narrow AI, for example have improved the work of farmers - in the global north as well as the global south -, logistic cooling chains, healthcare and energy management of households. Large companies already use so many AI applications that they struggle to keep track.

The rapid uptake and development of AI has also put the other side of the equation into sharper focus: the risks that come with the use of AI systems and the possible harm they might do. The concerns range from an AI system providing the wrong advice, “hallucination results” produced by ChatGPT, undermining democracy with fake news and deep fakes, to the destruction of humanity. These concerns cannot but be taken seriously, as confidence and trust are crucial for the acceptance and the success of any new technology.

To achieve a sufficient level of trust is a challenge in itself. The launch of ChatGPT has power-charged the debate about the pros and cons of in particular generative AI and raised the level of concern. ‘Founding fathers of AI’ have called for a pause in further development and CEO’s of AI developing

¹ Serentschy Advisory Services GmbH

² [Leaders Communique 01 en.pdf \(g7hiroshima.go.jp\)](#)

companies are calling for regulation. Proposals have been made to only allow further development of generative AI models in a safe, well-protected environment supervised by governments.

Against the background of these developments, the European Parliament has scrambled to include foundation models in its recent Report on the regulation of AI. This contains amendments³ to the original legislative proposal for an EU Artificial Intelligence Act (AIA) introduced by the European Commission in April 2021⁴. The Report by co-rapporteurs Brando Benifei (S&D) and Dragos Tudorache (Renew) is scheduled to be debated and voted on Tuesday the 13th of June 2023.

With the adoption of the Report, the position of the European Parliament has been decided. The legislative process on the EU AI Act will then enter the last stage⁵ - that of trilateral negotiations between the European Commission, the Parliament and the Council of Ministers (EU Member States). This is no small task. The Parliament proposes 771 amendments, some of which will require extensive discussion. Nevertheless, all three parties seem to be well aware of the global pressure to regulate – compare the interest in the 16th of May US Senate hearing – and the time pressure resulting from the European elections due in May 2024.

In case the EU AI Act will be adopted in time before these elections, its entry into force will only be after 24 months, i.e. at the earliest in Q3 2025 and at the latest in Q2 2026. Given the speed of innovation, the two EU Commissioners responsible for digital matters, Vice-President Margrethe Vestager and Commissioner Thierry Breton have recently proposed voluntary approaches to bridge the gap. Whether these approaches will bring about the hoped-for anticipation by stakeholders remains to be seen, as their thrust and approach differ.

With an adopted AI Act, the EU will be setting an important international marker. After its entry into force, companies that want to put AI systems on the EU market or in operation in the EU will have to comply with the requirements. The fines for non-compliance are high.

Conformity assessment and enforcement will be a tough challenge. Not only for developers, producers, distributors, importers and deployers (users). It will be a big task for the Commission itself, for the (EU Parliament proposed) AI Office, the Member States and their competent authorities. A key question is sufficient human expertise, qualitatively as well as quantitatively. This applies to authorities at the EU and national level, but also to the third party bodies ('notified bodies') that will be tasked with auditing and conformity assessment. Of course, human resources will be an issue for private sector stakeholders as well.

Comparing regulatory approaches

Efforts to set up a global system of AI governance have just started. Following the meeting in Hiroshima, a G7 working group is to come up with proposals⁶. UK Prime Minister Sunak was expected to suggest the UK as a suitable basis for a global AI Authority during his recent visit to Washington.

³ [PR_COD_1amCom \(europa.eu\)](#) (for foundation models see amendment 339 article 28b)

⁴ Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, Brussels, 21.4.2021 COM(2021) 206 final

⁵ In first reading

⁶ "In this respect, we task relevant ministers to establish the Hiroshima AI process, through a G7 working group, in an inclusive manner and in cooperation with the OECD and GPAI, for discussions on generative AI by the end of this year. These discussions could include topics such as governance, safeguard of intellectual property rights including copy rights, promotion of transparency, response to foreign information manipulation, including disinformation, and responsible utilization of these technologies" (G7 Communique, 20 May 2023)

Amongst the countries of the OECD, and in a broader sense those that respect a rules-based international order, there is agreement on the main principles for ethical and trustworthy AI. And in the Transatlantic context, the US, the UK and the EU share concepts like a risk-based approach and the importance of standards. Yet, there are important differences. While the EU is in the last stretch to adopt legislation that requires conformity assessment before market entry, the UK and the US favor – for the time being - a non-legislative approach that does not require pre-market approval. Within the same Transatlantic context Canada is on the way towards a legislative approach: the Artificial Intelligence and Data Act (AIDA). Based on similar principles, but contrary to the EU AI Act, the Canadian draft law leaves the actual regulation to provincial authorities.

Broadly spoken, all approaches in different jurisdictions are based on the same principles. They work towards regulation that aims to balance opportunities and benefits against risk and harm to people's health, security or fundamental rights. In its amendments, the European Parliament also adds democracy and the environment to the list.

In the US, this balancing act is visible in the Blueprint for an AI Bill of Rights⁷ (2022) on the one side and the National AI Initiative Act (2020) on the other side⁸. The latter aims to ensure US leadership in AI research and development and the use of trustworthy AI systems in the public and private sectors. Other purposes of the NAIA are preparing the workforce for the integration of AI systems 'across all sectors of the economy and society' and the coordination between government, academia and civil society in AI R&D. The Blueprint formulates five principles 'that should guide the design, use and deployment of automated systems to protect the American public in the age of AI'. These are: safe and effective systems, algorithmic discrimination protection, data privacy, notice and explanation and human alternatives, consideration and feedback (see annex 1).

The UK government in March published its 'Pro-innovation approach to AI regulation'⁹ The approach mentions as essential characteristics: pro-innovation, proportionate, trustworthy, adaptable, clear and collaborative. The non-statutory framework is guided by five principles: 1. safety, security and robustness, 2. appropriate transparency and explainability, 3. fairness, 4. accountability and governance and 5. contestability and redress. While the awareness of risk and harm is clearly visible in these principles, the ambitious set of objectives of the 'lean' UK approach focuses on opportunities and innovation: strong collaboration between government, regulators and business, empowerment of existing regulators (so no new, special AI regulating body), making responsible innovation easier, strengthening the UK's position as a global leader in AI, harness the UK's ability to drive growth and prosperity, increase public trust and ensure continuous adaptability (see annex 2)

The EU approach is flowing from the logic of the rules of the Internal Market on product safety and consumer protection. This is why it not only aims to balance the benefits and opportunities with the prevention of harm and the protection against risks, but it is based on the existing principles and mechanisms of conformity assessment before allowing a product to be put on the European market. The system of the EU AI Act relies on proven concepts of a division of labor between the EU level and national authorities for compliance and enforcement as well as for ensuring fair competition (see annex 3 for the 'what and how').

How to keep the balance?

Since the global availability of Large Language Models, the debate has tilted towards the negative side of the balance; towards risks and harm – even fear. In his recent FT article 'Regulating AI is a 4D

⁷ [Blueprint for an AI Bill of Rights | OSTP | The White House](#)

⁸ [National Artificial Intelligence Initiative Act of 2020 \(ai.gov\)](#)

⁹ [a pro-innovation approach to AI regulation print ready version.pdf \(publishing.service.gov.uk\)](#)

challenge'¹⁰ John Thornhill listed four challenges for AI regulation: discrimination (e.g. social scoring), disinformation (e.g. disrupting democracy with deep fakes), dislocation (of the labor market) and devastation (WW3, lethal autonomous weapons). While the first seems manageable with a prohibition, the others are more complicated from a regulatory point of view. Experience from nuclear non-proliferation treaties may serve as an example how effective such rules are and how difficult it is to enforce them.

This is only an illustrative selection of concerns and fears that powerful AI might get out of hand. As time progresses, we will discover more risks and possible harms. Thornhill rightly points to the fact that laws evolve incrementally whereas AI is developing exponentially. In regulating technology that ticks according to a Moore's clock on steroids, this is a continuous dilemma. But it is possible to put the extent of the dilemma into perspective if we consider how to deal with the different types of AI.

A common way to describe the different types of AI is the division between weak (or narrow) AI, strong (or generalized) AI and super (or conscious) AI. While narrow AI has been in use for considerable time in a very broad field of applications, generalized and generative AI have taken off exponentially more recently. Whether super or conscious AI is in existence falls outside the scope of this paper. But it is an important point on the horizon that might be closer than we are aware.

Using the EU proposed classification of low and high-risk AI systems, let us take a look at whether a balance between risks/harms on the one side and benefits/opportunities on the other can be found for narrow AI applications. For this, we look at three examples of use-cases of AI-for-a-purpose, each time in a low risk and a high-risk application.

1. A *smart home* is provided with an AI system to manage the consumption of electricity. These systems include a smart meter. The risk level can be considered low, as long as the system restricts itself to recommendations to the end user. Upgrading to an electricity management system that bases itself for example on the behavior of the consumer would mean that the risk level would increase. A consumer whose health depends on an electrical respiration device during sleep could be at risk if the AI wrongly interprets her going to bed as lying down on the living room sofa.
2. A *drone used by a farm* to collect data for an AI for crop prediction falls into the low risk category. But when the AI system is upgraded to allow autonomous decisions about the dissemination by the drone of pesticides and fertilizer, the AI would move towards a high-risk application.
3. A *healthcare platform* uses an AI system to process (anonymized) medical data to provide public and private organizations with better insights, for example an insurance company. As long as data regulation rules are respected, this can be considered a low risk activity. But if the platform also serves an AI driven alarm-function based on a symptoms-checker that allows external monitoring of a patient, things can move into high-risk territory. An elderly patient might have the habit of stumbling, and a sudden fall caused by a cardiac attack be misinterpreted, resulting in the failure to send an ambulance.

With narrow AI use cases like these, sustaining a balance between risks and benefits seems possible, even when the AI application is considered involving a higher risk. But there are already many of these systems in use and their number will further increase. While the EU regulatory approach would require a substantial amount of work to re-ensure compliance with every change of purpose, the UK and US approaches would likely be able to deal with the changes in a more flexible manner.

¹⁰ Harvard Business Review, 19 May 2023 [Who Is Going to Regulate AI? \(hbr.org\)](https://hbr.org)

Two elements seem important in this respect: the involvement of all stakeholders and continuous adaptation. The chances of keeping the balance between the risks and benefits of fast-moving technological innovation increase if all stakeholders contribute to a revolving “self-learning” regulatory mechanism. Such a mechanism is learning on the basis of continuous monitoring, based on collaboration of stakeholders in the public and private sector as well as consumers.

Introducing a ‘balancing’ timeline

Now let us look at the more complex part. Generative AI is difficult to grasp. John Thornhill describes its nature as ‘invisible, pervasive and having infinite use cases’ (FT 25 May 2023). What would be an acceptable level of risk if we consider an effective answer to challenges like disinformation and dislocation (e.g. of existing jobs) as the benefit – or better: *opportunity* - on the other side of the equation? The more we understand of the global challenges of the near future, the more it seems justified to focus on the benefits part of the balance and risks as a price we might be willing to pay.

Moving fast into unknown territory – some people think ‘God-like AI lies around the corner – might be done with more confidence if we move towards the acceptance of a balance between justified concerns about risk and harm and opportunities we cannot afford to miss for a certain price. Should we not be interested in the use of a strong AI system to solve societal issues resulting from climate change and migration, at the price that the same AI might create stress in the democratic system?

We might want to re-align our thinking about the benefits and opportunities of strong and super AI (generative AI). Assessing the price of the risks involved will be a collective task for all those involved in the creation, marketing, use, monitoring and managing of AI systems. It is an exercise democratic systems should be good at. An agile, revolving regulatory model could be a great asset to support this act of continuous balancing.

List of references

G7 Hiroshima Leaders’ Communique [Leaders Communique 01_en.pdf \(g7hiroshima.go.jp\)](#)

Report from the European Parliament Committee on the Internal Market and Consumer Protection and the Committee on Civil Liberties, Justice and Home Affairs on the EU AI Act 22 May 2023 [PR_COD_1amCom \(europa.eu\)](#)

Proposal for a Regulation of the European Parliament and the Council laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts, Brussels, 21.4.2021 COM (2021) 206 final

Blueprint for an AI Bill of Rights, The White House, October 2022 [Blueprint for an AI Bill of Rights | OSTP | The White House](#)

National Artificial Intelligence Initiative Act, Washinton 2020 [National Artificial Intelligence Initiative Act of 2020 \(ai.gov\)](#)

A Pro-Innovation Approach to AI Regulation, UK Government White Paper, March 2023 [a_pro-innovation_approach_to_ai_regulation_print_ready_version.pdf \(publishing.service.gov.uk\)](#)

‘Regulating AI is a 4D Challenge’ John Thornhill, Financial Times 25 May 2023

‘Who is going to regulate AI’, Harvard Business Review, 19 May 2023 [Who Is Going to Regulate AI? \(hbr.org\)](#)

Annex 1



Five principles that should guide the **design, use, and deployment** of automated systems to protect the American public in the age of artificial intelligence.

Safe and Effective Systems

You should be protected from unsafe or ineffective systems.

Algorithmic Discrimination Protections

You should not face discrimination by algorithms and systems should be used and designed in an equitable way.

Data Privacy

You should be protected from abusive data practices via built-in protections and you should have agency over how data about you is used.

Notice and Explanation

You should know that an automated system is being used and understand how and why it contributes to outcomes that impact you.

Human Alternatives, Consideration, and Fallback

You should be able to opt out, where appropriate, and have access to a person who can quickly consider and remedy problems you encounter.

Source: White House Office of Science and Technology Policy – White Paper



National Institute of Standards and Technology

National AI Initiative Act (2020)

The purposes of the Initiative are:

1. ensuring continued United States leadership in artificial intelligence research and development
2. leading the world in the development and use of trustworthy artificial intelligence systems in the public and private sectors;
3. preparing the present and future United States workforce for the integration of artificial intelligence systems across all sectors of the economy and society; and
4. coordinating ongoing artificial intelligence research, development, and demonstration activities among the civilian agencies, the Department of Defense, and the Intelligence Community to ensure that each informs the work of the others.

Annex 2

A pro-innovation approach to AI regulation

March 2023

“We have identified the essential characteristics of our regulatory regime: [pro-innovation](#), [proportionate](#), [trustworthy](#), [adaptable](#), [clear](#) and [collaborative](#). “

- [Pro-innovation](#): enabling rather than stifling responsible innovation.
- [Proportionate](#): avoiding unnecessary or disproportionate burdens for businesses and regulators.
- [Trustworthy](#): addressing real risks and fostering public trust in AI in order to promote and encourage its uptake.
- [Adaptable](#): enabling us to adapt quickly and effectively to keep pace with emergent opportunities and risks as AI technologies evolve.
- [Clear](#): making it easy for actors in the AI life cycle, including businesses using AI, to know what the rules are, who they apply to, who enforces them, and how to comply with them.
- [Collaborative](#): encouraging government, regulators, and industry to work together to facilitate AI innovation, build trust and ensure that the voice of the public is heard and considered.

A pro-innovation approach to AI regulation

March 2023

Five principles to guide and inform the responsible development and use of AI in all sectors of the economy

1. [Safety, security and robustness](#)
 2. [Appropriate transparency and explainability](#)
 3. [Fairness](#)
 4. [Accountability and governance](#)
 5. [Contestability and redress](#)
- [Non-statutory](#)
 - Relying on [collaboration](#) between government, regulators and business
 - empower [existing regulators](#) and promote coherence across the regulatory landscape.
 - [make responsible innovation easier](#).
 - [strengthen the UK's position as a global leader in AI](#),
 - [harness AI's ability to drive growth and prosperity](#)
 - [increase public trust in its use and application](#).
 - [agile and iterative approach, monitoring](#)
 - [learn from experience and continuously adapt](#) to develop the best possible regulatory regime.

Annex 3

EU AIA Recital nr 1 (EP-amended version)

what

The [purpose](#) of this Regulation is

1. to promote the uptake of human centric and trustworthy artificial intelligence and
2. to ensure a high level of protection of health, safety, fundamental rights, democracy and rule of law and the environment from harmful effects of artificial intelligence systems in the Union
3. while supporting innovation and
4. improving the functioning of the internal market.

how

This Regulation lays down a [uniform legal framework](#) in particular

1. for the development, the placing on the market, the putting into service and the use of artificial intelligence in conformity with Union values and
2. ensures the free movement of AI-based goods and services cross-border,
3. thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorised by this Regulation.