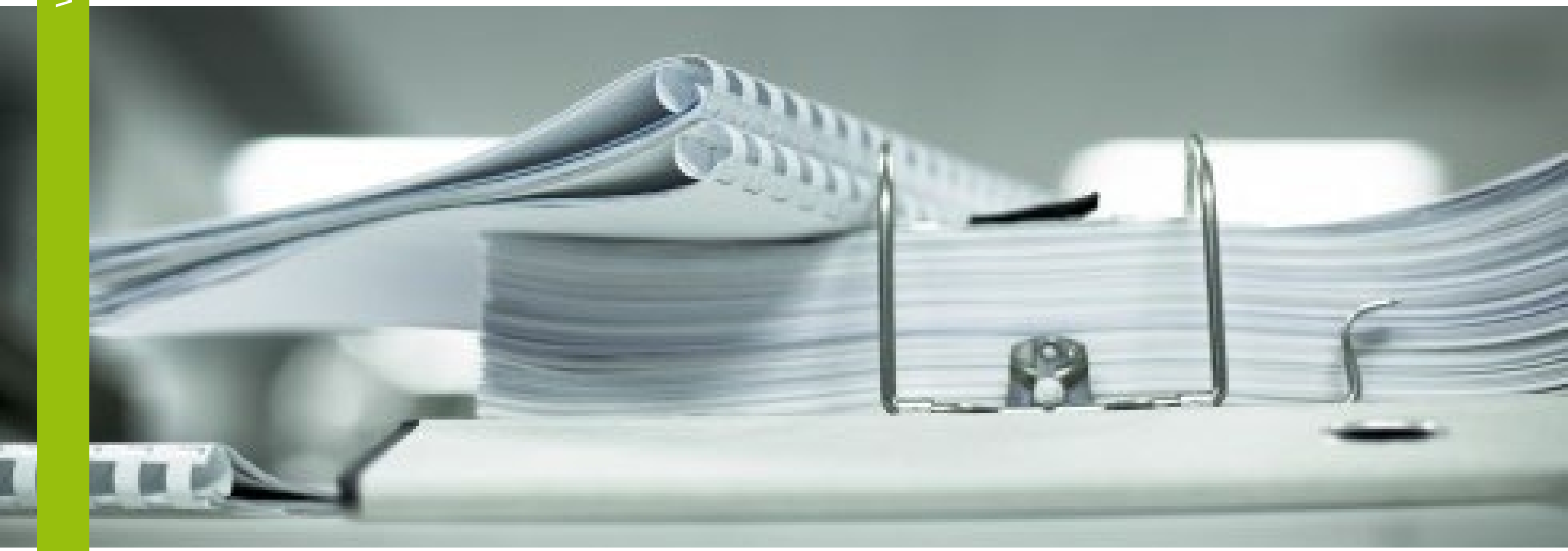
Policy Brief | 19.09.2023



The European Parliament’s amendments to the AI Act

A plausible approach for the “Ecological Alignment" of AI? [1]

// Peter Gailhofer (Öko-Institut e.V.), Martin Führ (Society for Institutional Analysis - sofia/Darmstadt University of Applied Sciences), Jens Gröger (Öko-Institut e.V.),
With the cooperation of Lars Nolle (Jade University of Applies Sciences) and Frederic Theodor Stahl (German Research Center for Artificial Intelligence)

According to the European Parliament’s (EP) amendments on the commission’s proposal for an Artificial Intelligence (AI) Act, adopted on 14 June 2023, a whole series of new, environment-related requirements for AI systems shall be integrated into the regulation’s framework. This paper outlines the amended framework and provides an assessment from an environmental point of view. The appraisal is guided by the political goal of a digital and environmental “twin transition” formulated in the EU’s “Green Deal”, according to which the legal framework must be designed in such a way that both transformative dynamics reinforce each other.

Against this backdrop, the paper draws the following conclusions: First and foremost, it is commendable that the parliament endeavours to tackle the pressing imperative of mitigating the substantial risks associated with the new technologies, encompassing not only "human dignity and personal autonomy" but also broader common good imperatives such as "social and environmental well-being”. This objective may be defined, borrowing from current international debates, as the “socio-ecological alignment of AI”. Secondly, the assessment shows that the proposed requirements for AI systems are technically feasible and appropriate in terms of their implementation effort for the regulation’s addressees; thus, the criteria of proportionality are met. Thirdly, however, the EP proposals do not fully remedy the shortcomings of the commission’s draft AI Act with regard to the environment. These flaws are due to specific gaps in the regulation’s new environmental provisions and, in a more fundamental manner, given the inadequate fit of its implementation and enforcement mechanisms to environmental risks. Thus, there remains a “delta” between the EP’s general objectives and the incentive situation provided by the amended legal

[1] This paper was prepared within the context of the research project "Regulatory framework for algorithmic decision-making systems under environmental law" funded by the German Environment Agency.

framework: Prominently, the scope of application of the systems to be included in the risk-based regulation is not adequately tailored to environmental issues. Fourthly, these shortcomings are partly mitigated by the fact that the draft regulation displays features of a learning system. The analysis suggests that these reflexive elements should be enhanced in order to bring environmental aspects more strongly into play.

Key policy recommendations

- The requirements provided for in the parliamentary amendments to identify, assess and mitigate environmental risks and impacts are reasonable and feasible. This holds, specifically, for the provisions to quantify, log and minimize energy consumption and resource depletion: Methods and technical means to specify these demands are already established or in sight. The proposed provisions should therefore be adopted as widely as possible in the trilogue.
- However, limiting most of these requirements in the proposal to "high-risk" applications is problematic. From an ecological perspective, the scope of relevant rules should be broadened to include ecologically sensitive areas or use-cases of AI systems, e.g. for system steering industrial installations under the Seveso or the Industrial Emissions Directive as well as water facilities under the Water Framework Directive.
- The “general principles” applicable to all AI systems formulated in Art. 4a of the proposal, which stipulate, legally speaking, dynamic and permanent “general obligations” for all relevant systems and actors, highlight the socio-ecological turn in the Parliament’s regulatory approach. They include commitments to democracy and “social and environmental well-being”. However, it is important to note that these principles for a comprehensive alignment of AI systems with socio-ecological goals – beyond concrete obligations that again only apply to specific high-risk systems – are not backed up by binding substantive and procedural requirements. Therefore, they do not provide clear incentives.
- The parliamentary proposal contains a number of mechanisms for “regulatory learning” that could help to broaden the scope of application and reflect new insights into environmental impacts. In order to comply with the general principles of the proposal on the socio-ecological alignment of AI systems, these learning mechanisms should be more consistently geared towards countering potentially serious emergent environmental risks. This could happen, for example, by extending transparency rules and by introducing collective rights of action and data access rights for civil society and the scientific community.

1 Context: Environmental risks of AI systems

The current trilogue negotiations on the AI Act come at a time when there is much debate about ethical and societal risks of AI. Increasing calls for regulation are prominently justified by catastrophic scenarios in which large parts of humanity are wiped out by flawed or malicious deployment of AI, or even by a "general-purpose", malevolent AI itself.[2] Such dire warnings might have their merits. But, what if more plausible existential threats do not lie primarily in AI-generated weapons or epidemics, but in more subtle and elusive dynamics through which AI systems fuel climate catastrophe and environmental degradation?

Environmental risks of AI systems have come increasingly into the spotlight. The focus of attention here lies on the CO_2 emissions of some systems: in particular, large AI models require huge amounts of energy for training and operation. If, as is widely predicted, such and similar systems become prevalent to a similar extent as PCs or smartphones, the use of resources for the necessary hardware may equally have massive environmental effects. The adoption of extremely energy- and resource-intensive technologies across society is likely to further exacerbate the already daunting challenges facing the necessary green transformation of the economy. It is little wonder then, that the media and sometimes also regulatory[3] debate concentrates on such so-called **direct environmental effects** of AI systems.

Less noticed, but equally important, are AI's **indirect environmental effects**, a facet of the much-discussed "alignment problem". This problem centres on the challenge of ensuring that intelligent systems don't pursue and – potentially very effectively – implement goals that conflict with fundamental values or human rights. Misalignment can arise from user or developer intentions, biased data, design errors, or unforeseen user feedback.[4] While the focus of the general alignment debate lies on more salient existential risks or, for example, immediate dangers of discrimination or manipulation, more and more research describes harmful impacts of AI systems, resulting from poor alignment with environmental goals and values: For instance, non-aligned AI in industrial applications can lead to increased energy or resource consumption due to prioritizing cost efficiency over sustainability.[5] In precision agriculture, misalignment can result in excessive nitrogen fertilizer use, driven by poor incentives, inadequate data, or unexpected user behaviour.[6] Examples abound: self-driving cars that no longer

---

[2] See, for example, Shevlane et.al (2023): Model evaluation for extreme risks, https://arxiv.org/abs/2305.15324, or https://www.alignmentforum.org/posts/MtDmnSpPHDvLr7CdM/catastrophic-risks-from-ai-2-malicious-use.

[3] Hacker (2023): Sustainable AI Regulation (June 1, 2023). Available at SSRN: https://ssrn.com/abstract=4467684.

[4] This paper employs a broad concept of alignment, as opposed to a concept that focuses on the technical difficulty of ensuring that the results of AI systems match the goals or preferences of their operators or users and thus do not reflect on the issue whether these goals or preferences are aligned with societal/ethical values, cf. Gabriel, I. (2020): Artificial Intelligence, Values, and Alignment. *Minds & Machines* 30, 411–437. https://doi.org/10.1007/s11023-020-09539-2.

[5] Carlson & Sakao (2020): Procedia CIRP 90, 174.

[6] Galaz et.al. (2021): Artificial intelligence, systemic risks, and sustainability, Technology in Society 67.

need a parking space because they circle the block while their owners are busy shopping;[7] "persuasive" shopping assistants driving their users to consume ever more and faster;[8] or chatbots flooding social networks with data without human involvement. Many examples illustrate that such harmful effects can often be viewed as the flip side of AI systems that, at first glance, seem well aligned with environmental goals.[9] If oriented towards unsustainable goals, however, they "[speed] up natural resource extraction, [supercharge] unsustainable consumption, and [deepen] the ecological shadows of global supply chains".[10]

AI alignment in general faces complex and partly unsolved challenges. *The ecological alignment problem*[11] adds an additional layer of complexity to these challenges. Environmental risks often stem from seemingly harmless applications that, when used on a large scale, lead to serious impacts. Unlike many AI-related risks that directly affect individuals, environmental risks often lack clear voices to raise concerns and draw attention. The challenges posed by the "black-box" nature of algorithms, where AI systems operate opaquely, are compounded by the opacity of environmental issues. It's often challenging to definitively attribute harmful consequences to environmental and climate impacts, and they may only become evident over time or in conjunction with other factors. Addressing the environmental risks associated with AI thus may require addressing an exacerbated "black box". This is particularly concerning because ecological alignment ex-post may prove difficult, especially as these new systems become deeply integrated into everyday life, industries, businesses, and government administration, heightening the risk of a "lock-in effect", where numerous dependencies make it harder to shift society's direction. [12] The convergence of these factors underscores the existential peril discussed earlier: subtle, often imperceptible misalignments within often widely utilized everyday systems may trigger a rapid escalation of the destructive forces surrounding climate change and environmental degradation, with potentially irreversible consequences.

The EU recognizes that achieving the ambitious goal of a green transition is incompatible with conflicting technological dynamics and has formulated the political goal of

---

7 Millard-Ball (2019): The autonomous vehicle parking problem, Transport Policy (Vol. 75), 99-108.

8 Cf. Sanderson (2023): Personalized Content Delivery with Generative AI: A New Frontier in Marketing, available online at: https://www.progress.com/blogs/personalized-content-delivery-generative-ai-new-frontier-marketing.

9 Cf., for example, Gailhofer et.al. (2021): The Role of Artificial Intelligence in the European Green Deal, Study for the special committee on Artificial Intelligence in a Digital Age (AIDA), Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg, available online at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/662906/IPOL_STU(2021)662906_EN.pdf.

10 Dauvergne (2022): Is artificial intelligence greening global supply chains? Exposing the political economy of environmental costs, Review of International Political Economy,29:3, 696-718.

11 Christian (2020): The Alignment Problem: Machine Learning and Human Values.

12 Robbins & van Wynsberghe (2022): Our New Artificial Intelligence Infrastructure: Becoming Locked into an Unsustainable Future. Sustainability 2022, 14, 4829.

a digital and ecological "twin transition",[13] as outlined in the Green Deal.[14] This political goal, however, can only be attained if the ecological alignment issue is adequately addressed within the regulatory framework.

2 The "socio-ecological turn" of the EU Parliament's amendments to the AI Act

The European Commission published its proposal for an AI Regulation in 2021 (COM(2021) 206 final).[15] It aims at ensuring safety and trustworthiness of AI, at establishing legal certainty to facilitate investment and innovation in the EU single market, and effectively govern potentially serious risks of AI systems, not least given their opacity, complexity, and partly autonomous behaviour.[16] In addition to specific prohibitions of particularly dangerous systems, the risk-based framework contains primarily rules on the precautionary management and transparency of so-called "high-risk" systems, as well as some requirements for systems that are considered less risky. Nonetheless, the commission's draft conspicuously lacks substantial consideration for ecological alignment.[17]

It is to the European Parliament's credit that it is seeking to fill this gap with its negotiation mandate of 14 June 2023 (hereinafter: EP Proposal or Proposed Amendments).[18] The partial shift of the parliament to broader regulatory goals, which include environmental concerns, is evident in the very first recital of the parliament's proposal. Rather than being primarily concerned with the internal market, it is now focused on the "uptake of human-centric and trustworthy" AI. The commission's secondary objective of a "high level of protection" has now moved to the first place. In particular, the parliament aligns this goal at the very beginning of the regulation with the – from the authors' point of view fundamental – goods of "democracy and rule of law and the environment". The intention to introduce an ecological dimension to the regulation is also set out in Article 1 of the parliamentary proposal as a regulatory objective.

This "environmental shift" is reflected throughout the entire proposal and is approached in different ways: First of all, the parliament proposes specific requirements

13 See, for example, the JRC "Strategic Foresight Report Twinning the green and digital transitions in the new geopolitical context", available online at: https://joint-research-centre.ec.europa.eu/jrc-news-and-updates/twin-green-digital-transition-how-sustainable-digital-technologies-could-enable-carbon-neutral-eu-2022-06-29_en.

14 European Commission (2019): The European Green Deal, COM(2019) 640.

15 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, COM/2021/206 final, online available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206.

16 See the commission's proposal for an AI Act of 21.04.2021, COM(2021) 206 final, Explanatory memorandum, 2-3.

17 See Pagallo, Sciolla & Durante (2022): The environmental challenges of AI in EU law: lessons learned from the Artificial Intelligence Act (AIA) with its drawbacks, Transforming Government: People, Process and Policy, Vol. 16 No. 3, pp. 359-376.

18 Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD)), online available at: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.

and obligations for high-risk systems and foundation models,[19] which are quite clearly focused on minimizing direct environmental effects. Second, the proposal integrates foreseeable and emerging environmental effects into the existing requirements and obligations for the identification, assessment and mitigation of risks, specifically those relating to high-risk systems. This means that indirect environmental effects also come into the focus of the AI Act (see section 2.1.). Third, some environmental requirements are set in a horizontal manner, i.e. as rules or principles applicable to all systems, irrespective of particular application areas (see section 2.2.).

2.1 Obligations for high-risk systems and foundation models

In the former sense, a new requirement for high-risk systems is proposed, which stipulates that, by design, logging functions must be in place to enable the recording of energy consumption, the measurement or calculation of resource consumption and the environmental impact of the high-risk AI system during all phases of the system’s life cycle, Article 12 para. 2 a. Comparable obligations are specified for the providers of foundation models: The systems must have the “capabilities to enable the measurement and logging of the consumption of energy and resources, and, if feasible, other environmental impact, over their entire lifecycle”, Article 28 b para. 2 (d). The same article also lays down the obligation to design the systems in such a way as to reduce energy consumption, resource use and waste and to ensure energy efficiency and overall efficiency of the systems.

In the latter sense, new references to environmental risks are being integrated into many of the existing requirements for high-risk systems and into obligations of critical actors in high-risk system’s lifecycle.

- Prominently, foreseeable and emergent environmental risks shall be identified, assessed and mitigated, for example, by means of suitable testing procedures, as part of a risk management system and continuously and iteratively executed throughout the system´s entire lifecycle, Article 9 para. 2 a.
- The results of the analysis and assessment must be documented, and systems must be designed in a way to ensure transparency regarding potential risks. Intelligible instructions shall be attached, which enable providers and users to understand the system`s functioning and, importantly, also inform about circumstances related to their use and misuse which could lead to environmental risks, Article 13 para. 3 b (iii).
- The systems must enable, by design, e.g. by appropriate interface tools, human oversight to mitigate or prevent environmental risks from the use of the systems, Article 14 para. 2.

Critical actors in the lifecycle of the systems – providers, operators, deployers – are charged with differentiated obligations to ensure compliance with the specifications for the systems, e.g. implementing a quality management system, duties to ensure conformity assessment procedures etc.

[19] Foundation models are systems trained on a particularly broad data basis which, like the current large language models, can be adapted for a large number of specialised applications (see Article 3 para. 1 (1c)).

- For foundation models, Art. 28 b para. 2. (a) establishes the obligation of providers to demonstrate that precautionary measures have been taken to identify, reduce and mitigate risks to the environment (through testing, design and analysis). This obligation applies "prior and throughout the development" and is to be carried out "with appropriate methods such as with the involvement of independent experts, as well as the documentation of remaining non-mitigable risks after development".
- A "fundamental rights impact assessment" has to be carried out by deployers of high-risk systems in the specific use-context and encompass "the reasonably fore-seeable adverse impact of the use of the system on the environment" according to Art. 29 a (g).

2.2 Horizontal requirements addressing environmental effects

The EP proposal also contains horizontal, i.e. general, environmental requirements that apply to all AI systems. Article 4a para. 1 of the proposal, which can be understood as a kind of socio-ecological fundamental obligation aiming at integrating public interest requirements into the design process of AI system, ranging from "human dignity and personal autonomy" (a) up to "social and environmental well-being" (f): This "means that AI systems shall be developed and used in a sustainable and environmentally friendly manner as well as in a way to benefit all human beings, while monitoring and assessing the long-term impacts on the individual, society and democracy". These fundamental obligations shall be concretised, for example, through voluntary codes of conduct, Article 69.[20]

3 Assessment: Provisions on environmental risks sensible, feasible but gaps remain

The EP's suggestions to address the challenges of ecological alignment are, in principle, sound and pragmatic.

3.1 Plausible instruments to assess, evaluate and mitigate environmental risks

This is true, first of all, for the newly proposed requirements for measuring, logging and reducing the resource and energy consumption (**direct environmental effects**) of AI systems themselves. The quantification of energy consumption in the production and operation of AI systems depends on variables at different stages of the life cycle (e.g. varying energy requirements of different models and methods, (cloud) infrastructure and energy source used, type and amount of data used).[21] The question of which methods and standards should be applied to deal with such factors is left to the sub-

[20] Furthermore, Art. 82 b of the EP proposal empowers the commission doe develop "guidelines on the practical implementation" of the regulation. In practical terms, these guidelines might steer the behaviour of the AI providers to a high extent. This effect can be seen in the implementation of the chemicals regulation REACH; cf. https://echa.europa.eu/support/guidance. However, these guidelines again mainly concern the implementation of the requirements for high-risk systems and foundation models.

[21] See, regarding the quantification of CO_2 emissions: Cowls et.al (2021): The AI gambit: leveraging artificial intelligence to combat climate change—opportunities, challenges, and recommendations, AI & Society 2021.

sequent development of specifications by the commission, to the adoption of implementing acts and by reference to concretising standards (also yet to be developed), Article 41 para. 1 c.

However, there is little to suggest that too much would be demanded. The providers of AI face the same challenge as all other providers of digital services. In the future, they will have to make transparent to their customers and regulatory authorities which environmental impacts their services are associated with.[22] For example method "Green Cloud Computing"[23] demonstrates how corresponding environmental indicators can be calculated. In accordance with accepted standards for life cycle assessment the method calculates greenhouse gas emissions, energy consumption, resource depletion and water usage of digital services (e.g., video streaming, online storage, desktop virtualization). The same method can also be used to calculate the environmental impacts of artificial intelligence applications and assign the overall results to individual use cases. Such calculation methods should already be applied voluntarily during the development process of AI applications. However, at the latest when AI applications of a certain scale are published and used commercially, they should make their environmental impacts transparent to the users on a mandatory basis. Conceivable dimensions for the obligation are, for example, based on the continuous electrical power demand of the information technology used (e.g. from 10 kW), or the annual turnover achieved with it (e.g. from 100,000 EUR).

Energy consumption, even of individual applications, can be visualised and made available by existing software tools. Large cloud providers offer their customers precise information about the energy consumption and related CO_2 emissions of the requested cloud services.[24] Software tools are available that are supposed to calculate energy consumption and related CO_2 emissions while programming.[25] The feasibility, in principle, of measurement and transparency regarding energy consumption is also demonstrated by the model card of the open-source model "Bloom".[26]

Energy consumption in the application phase could be evaluated ex ante by typifying the application (instantiation) of the models, its data-intensity and by looking at the expected frequency of usage. The factual frequency of usage can be assessed by live-monitoring the systems. Given that the performance of the systems must be systematically monitored after they have been placed on the market according to Article 61 of the proposal, the post-market monitoring of direct environmental effects might be implemented with little additional effort. The obligation to design systems as energy- and resource-efficient as possible also seems feasible. Research argues that

---

22 Bilsen et al. (2020): Study on Greening Cloud Computing and Electronic Communications Services and Networks: Towards Climate Neutrality by 2050, https://digital-strategy.ec.europa.eu/en/library/study-greening-cloud-computing-and-electronic-communications-services-and-networks-towards-climate.

23 Gröger, Liu, Stobbe, Druschke, Richter (2021): Green Cloud Computing. Lifecycle-based data collection on the environmental impacts of cloud computing. https://www.umweltbundesamt.de/publikationen/green-cloud-computing.

24 https://cloud.google.com/carbon-footprint/docs/methodology?hl=de.

25 https://mlco2.github.io/codecarbon/methodology.html.

26 https://huggingface.co/bigscience/bloom; also see, for example, Anthony, Kanding & Selvan (2020): Carbontracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models, online available at: https://arxiv.org/pdf/2007.03051.pdf.

choices of the deep neural network, data centre and processors to be used can reduce the carbon footprint by a factor of 100 to 1000.[27] Its critical to exploit such potentials for efficiency.

It is encouraging that the Parliament's proposed amendments also focus risks for (negative) **indirect environmental effects**. The provisions regarding iterative measures to identify, assess and mitigate environmental risks over the system's entire life cycle necessarily look at risks of the intended application, and at unintended effects, which, for example, may result from interactions with human users or other AI systems. The regulatory approach of the AI Act relies on specifications for the systems themselves, and organisational and procedural rules that are to be autonomously implemented by different actors in the lifecycle. The **self-regulation of providers**, deployers and operators is evaluated by means of conformity assessments and against the benchmark of partly private standards and controlled and sanctioned by public authorities. The implementation and enforcement of the AI Act also relies on external, e.g. scientific, inputs to take account of technical developments, emergent risks and appropriate approaches to solutions. It contains a series of rules, many of which introduced by the parliament, which may justify its classification as a "reflexive regulation", i.e. one that continuously learns from practice and extra-legal expertise.[28] Similar regimes have already proven feasible elsewhere in environmental law to cope with unknown, complex and emergent risks. There is no reason, why this should not be the case, in principle, with respect to AI-systems.

3.2 Remaining deficits from an environmental point of view

If the proposed amendment falls short of what is necessary, this is therefore not due to a general lack of "fit" or practicability of its instruments for the integration of ecological risks.

To be sure, environmental effects are not taken into account at some points in the risk management process where this would have seemed obvious. For example, it is hardly comprehensible that the rules on **data governance**[29] should remain without reference to environmental risks. Harmful data biases and self-reinforcing feedback loops can be considered as relevant causes of indirect environmental effects of AI systems.[30] The respective provisions of the parliamentary proposal (Art. 10 para. 2 f.) could easily integrate such risks of specific ecological data biases.

In addition to such rather specific shortcomings, **structural deficits** remain.

Firstly, the **implementation and enforcement mechanisms** of the regulation remain insufficient to meet the challenges to detect and effectively mitigate environmental risks. It has been aptly put that approaches of regulated self-regulation can only be

---

27 Patterson et. al. (2021): Carbon Emissions and Large Neural Network Training, arXiv:2104.10350, online available at: https://arxiv.org/ftp/arxiv/papers/2104/2104.10350.pdf. Also cf. OECD (2022): Measuring the environmental impact of Artificial Intelligence compute and applications, OECD Digital Economy Papers No. 341.

28 Gaines (2002): Reflexive Law as a Legal Paradigm for Sustainable Development, 10 Buff. Envtl. L.J. 1 (2002-2003).

29 See Art. 10 (2) f.) Parliamentary Proposal.

30 See Gailhofer & Franke (2021): Datenregulierung als sozial-ökologische Weichenstellung, Zeitschrift für Umweltrecht (ZUR 2021, 532); Franke & Gailhofer (2021): Data Regulation and Data Governance for Sustainable Smart Cities, Front. Sustain. Cities Vol.3, with further references.

as effective as the "shadow of law" and the related behavioural incentives behind it.[31] In order to deal with environmental risks and their characteristics (no directly affected people, cumulative, dispersed, time-delayed and indirect effects), broad, decentralised mechanisms would be needed to detect violations of environmental obligations and provide effective mechanisms of subsequent charging.[32] Transparency as well as rights of associations to initiate proceedings and other rights for non-governmental actors could contribute to the detection and mitigation of environmental risks.[33] The AI Acts provisions however remain very much concentrated on monitoring and auditing by public authorities, which may not have sufficient information about emergent environmental risks and generally have to deal with staff constraints. The (new) complaints mechanisms proposed by Parliament are once again very much tailored to "traditional" AI risks with little relevance to environmental issues.[34] A general allocation of the burden of proof, which could help to reduce information deficits of the authorities on environmental risks, similar to chemicals legislation, for example, is not conceivable. Only for "foundation models" (see section 2.1) Art. 28 b stipulates the duty for the provider to demonstrate that certain risk mitigation measures have been applied and to document the remaining non-mitigable risks.

Secondly, **the substantive scope of the regulation remains almost exclusively focused on systems that have little relevance to the environmental risks outlined above**. The criteria for the classification of high-risk systems in Annexes II and III, which determine the applicability of the vast majority of the "hard" rules of the regulation, have remained virtually unchanged. As a result, environmental risks must now be analysed, assessed, and mitigated, but primarily for systems that are supposed to pose a significant risk not because of their environmental relevance, but for entirely different reasons.[35] A regulation that specifically targets environmental risks would, of course, look different: Annexes II and III could have been amended to categorize AI systems as high-risk, which are particularly prone to environmental risks. Definitions of AI applications where those risks typically occur can be found in the Seveso Directive (2012/18/EU) and the Industrial Emissions Directive (2010/75/EU). Moreover, malfunctions of AI systems steering water treatments and drinking water facilities, as defined in the Water Framework Directive (2000/60/EC), bear typically high risks for human health and the environment.

---

31 Spieker genannt Döhmann & Towfigh (2023): Automatisch benachteiligt, Legal opinion commissioned by the Federal Anti-Discrimination Agency, 67.

32 Solano, Martin, de Souza & Taylor (2022): Governing data and artificial intelligence for all, Study – Panel for the Future of Science and Technology, EPRS, 60.

33 Pagallo, Sciolla & Durante (2022): The environmental challenges of AI in EU law: lessons learned from the Artificial Intelligence Act (AIA) with its drawbacks, Transforming Government: People, Process and Policy, Vol. 16 No. 3, pp. 359-376. For the criticism of the TÜV Bundesverband that the notified conformity assessment bodies do not have sufficient access to data in order to fulfil the requirements of the AI Act, see TÜV Bundesverband (July 2023), Position Paper Recommendations for the AI Act trilogue negotiations, 3.

34 See Art. 68a, 68b, 68d Parliamentary Proposal.

35 Cf. in particular Article 6 para. 2 of the proposal – even significant environmental risks can only justify classification as high-risk if they are identified for specific systems (Annex III No. 2). An exception to the lack of environmental focus is to be made with regard to the rules concerning the direct environmental effects of foundation models: these are legally defined in Article 3 in such a way that they are trained on broad data at scale (which causes a high energy input); precisely these systems are heavily criticized because of their energy consumption.

3.3 The Proposal's mechanisms for regulatory learning

These gaps are **partly mitigated** by the parliamentary proposal's "reflexive", i.e. adaptive approach. Importantly, in contrast to the commission's proposal,[36] areas or use-cases of high-risk AI systems can be added to the regulation, see Article 7 of the proposal.[37] Delegated acts may add or modify use-cases and areas of high-risk systems in Annex III, inter alia, if they are found to pose adverse impacts on the environment. As such amendments presuppose sufficient information regarding currently unknown environmental risks or impacts, the parliamentary proposal contains a number of provisions that are suitable, in principle, for generating this knowledge: Already the commission proposal included provisions on setting up regulatory sandboxes; the EP enhanced this approach substantially (Article 53, 53 a para. 2 h) and 54). A similar function serve the provisions on promoting AI research and development in support of socially and environmentally beneficial outcomes (Article 54 a). In this respect, sup-para 2 provides that member states "shall support civil society and social stakeholders to lead or participate in such projects". Furthermore, the text of the regulation itself is subject to continuous learning cycles, including participatory procedures for the evaluation and amendment of the regulation (cf. Article 84 para. 7 a).

One of the probably most relevant amendments in the EP proposal are somewhat hidden in Art. 56. Whilst the commission foresees a "Board" with the function to "provide advice and assistance to the commission" the EP wants to establish an "AI Office" as an "independent body of the Union. It shall have legal personality." With this shift of the organisation nature the EP establishes a new "player" in the implementation of the regulation. The office is structured (Art. 57 a and 58) in the same manner as a (regulative) agency (e.g., the European Chemicals Agency, ECHA, under REACH) and has competences that could not least strengthen the regulation's adaptability and dynamic implementation. It comprises according to Art. 56 a (a) "a management board, including a chair, (b) a secretariat managed by an executive director; (c) an advisory forum". Art. 56 b lists a long number of tasks, including the right to "h) examine, on its own [the secretariats] initiative or upon the request of its management board or the commission, questions relating to the implementation of this regulation and to issue opinions, recommendations or written contributions". This task is not limited; the examples ("including") are given under i) to vi) show that all kinds of relevant developments might be addressed. According to Art. 56 para. 2. the AI Office "shall be adequately funded and staffed for the purpose of performing its tasks pursuant to this Regulation." The trilogue will show in how far the EP can pursue this approach and secure a funding level that allows to establish the appropriate level of expertise in this independent body. It is hard to overestimate the fact that the Office is not directly under the control of the European Commission. It thus escapes the antagonisms and the resulting blockades that can often be observed between the individual Directorates General. The Office opens up the potential to bring an effective voice to the implementation of the regulation. In view of the proposed structure, it would then also only be a small step to a "mature" agency with independent regulatory competences.[38]

[36] According to the commission's proposal, high-risk systems can only be integrated if they, in addition to other criteria, can already be assigned to one of the areas of Annex III.
[37] See rec. (27) of the parliament's proposal.
[38] For the analysis of the similar structure under REACH see Führ (2014): Vom Wesen Europäischer Agenturen, in: Ewer/Ramsauer/Reese/Rubel (Hrsg.), Methodik - Ordnung -

The learning environment outlined above in brief strokes may help to identify and evaluate emergent environmental risks. Some of the mechanisms for regulatory control and enforcement might also help to address environmental risks of non-high-risk systems. For example, Article 65 allows competent authorities to classify such systems as presenting a risk, (amongst other things) if their environmental effects are disproportionate to their purpose. As consequence they shall evaluate and can eventually even require to withdraw the systems from the market.

For the purpose of recognising and preventing the risks of ecologically misaligned systems, however, these learning mechanisms would also have to be designed more stringently. Again, the effective levers to set impulses for regulatory learning, e.g. to gain deeper insights that could reveal ecological biases in training data or potentially harmful objective functions, lie with the authorities (see, e.g. Article 64 para. 1, Article 65).[39] Again, given the typical features of environmental risks (see above), the AI Act should go further here. It should empower actors who typically have an interest in uncovering environmental risks, for example, by regulating access to relevant data for vetted researchers, as provided for in the Digital Services Act.[40]

Umwelt (Festschrift für Hans-Joachim Koch), Berlin 2014, S. 229 - 252 (Duncker & Humblot).

39 The information to be published in a public EU database according to Articles 51, 60 and Annex VIII Parliamentary Proposal seem rather not sufficient for such analysis.

40 More generally, data regulation offers strong and largely untapped levers for an ecological alignment of AI systems, see Gailhofer & Franke (2021): Datenregulierung als sozial-ökologische Weichenstellung, Zeitschrift für Umweltrecht (ZUR 2021, 532); Franke & Gailhofer (2021): Data Regulation and Data Governance for Sustainable Smart Cities, Front. Sustain. Cities Vol.3; and Solano, Martin, de Souza & Taylor (2022): Governing data and artificial intelligence for all (2022): Study – Panel for the Future of Science and Technology, EPRS.

Conclusion

The European Parliament responds to the growing evidence on the significant environmental risks of AI systems and proposes a range of environmental and climate protection-related amendments to the AI Act. These proposals are both feasible and technically achievable, particularly in terms of measuring and mitigating energy and resource consumption in these systems. If the EU wants to achieve its environmental and digital policy goals of a "twin transition", these rules should be fully enacted.

The "general principles" applicable to all AI systems formulated in Art. 4 a stipulating, legally speaking, dynamic and permanent "General Basic Obligations" for all relevant actors are at the heart of the Parliament's socio-ecological turn. They include democracy and "social and environmental well-being". In essence, they establish a duty to conduct a socio-ecologic impact assessment at a very early stage of the design process but also during the use phase of AI systems. However, it is important to note that these principles for a broad alignment of AI-systems with socio-ecological goals are not backed up by binding requirements or clear incentives. Beyond concretizing obligations in the proposal that, again, only apply to high-risk systems, they rather are translated into recommendations and non-binding standards.

Other elements of the parliamentary proposal also reflect notable gaps from an environmental standpoint. For instance, the rules on data governance do not address data-related environmental risks. Furthermore, the regulation's scope, determining which systems are considered high-risk and subject to regulation, lacks an environmental perspective. As a result, environmental concerns and risks are primarily addressed in systems regulated for entirely different reasons. By contrast, applications that may be particularly risky from an ecological point of view may remain outside the scope of application.

The parliament's reluctance can presumably be explained by the fact that it wants to avoid the impression of a regulation that inhibits innovation. There may be concern that broad-based regulations could stifle the potential of the new technologies to promote economic growth and, ultimately, welfare. In view of the still existing knowledge-gaps regarding the severe consequences of ecologically non-aligned systems, comprehensive, horizontal obligations may seem inappropriate.

A sensible consequence of such concerns could involve leaning even more on the reflexive elements of the regulation, i.e. to promote the generation of regulatory knowledge and to establish the foundation for quickly adapting the regulatory framework as needed.

From an environmental perspective, it is therefore essential, that the regulatory framework offers channels and arenas for science and civil society to raise environmental concerns, e.g. through enhanced involvement into the learning mechanisms but also in terms of collective action rights of associations and extended complaint rights vis-à-vis the competent authorities, or comprehensive scientific data access vis-à-vis the providers or operators of AI systems. The latter, in particular, would align with the approaches of the Digital Services Act, the Data Act and the Data Governance Act and could have a facilitating rather than inhibiting effect with regard to AI innovations. Greater available knowledge on the challenges and potentials of ecological alignment would not least reduce the risk arising from the insufficient use of AI for ecological purposes.[41]

Öko-Institut e.V | Freiburg | Darmstadt | Berlin

The Oeko-Institut is one of Europe’s leading independent research and consultancy organisations working for a sustainable future. Since its establishment in 1977, it has been laying the groundwork and devising strategies to realise the vision of sustainable development at global, national and local level. The Oeko-Institut has offices in Freiburg, Darmstadt and Berlin.
www.oeko.de | info@oeko.de

Contact

Dr. Peter Gailhofer | +49 30 405085-352 | p.gailhofer@oeko.de

41 Pagallo, Sciolla & Durante (2022): The environmental challenges of AI in EU law: lessons learned from the Artificial Intelligence Act (AIA) with its drawbacks, Transforming Government: People, Process and Policy, Vol. 16 No. 3, pp. 359-376.