The Speinshart Recommendations on Generative AI and the EU AI Act

SEPTEMBER 2023

INTRODUCTION

Recent technological advancements have brought about a significant qualitative transformation in the realm of technology, particularly in the domain of generative AI. The incorporation of specific provisions concerning foundation models and generative AI into the proposed EU AI Act necessitates the compilation of policy−relevant considerations for policy actors and stakeholders, amid the ongoing global discussions on AI regulation and its impacts on people and communities.

The well−recognized challenge of regulatory lag, a consequence of rapid technological progress, is juxtaposed with the risk of overly specific regulations that may not remain adaptable over time. Thus, while expressing our broad appreciation for the EU's initiative, as reflected in the latest text ratified by the European Parliament on 14th June 2023, our discourse adopts a principled approach rather than a rigidly rule−based one. This facilitates an in−depth exploration of the socio−technical dynamics of AI systems within diverse sectoral contexts and allows a discussion of best practices ahead of the likely adoption of the EU AI Act in 2024 and with transition period till mid of 2026.

The Speinshart Symposia convened an interdisciplinary conference from 26th to 28th June 2023, focusing on the challenges and opportunities pertaining to Generative AI and the EU AI Act at the international science retreat center, Speinshart. In the framework of the Hightech Agenda Bavaria distinguished scholars from computer science, philosophy, law, public policy, alongside representatives from civil society and industry, convened to discuss the far−reaching implication of generative AI across the AI ecosystem.

Collectively, 22 action−oriented recommendations were formulated, aiming to provide meaningful guidance and discussion points for both generative AI and policymaking. To this end, the following framework was employed to elucidate and analyze the communication and interactions among multiple stakeholders and spheres within the AI ecosystem:

● Communication between regulators and actors within the regulated value chain (yellow).
● Communication among actors within the value chain (blue).
● Communication between affected societal stakeholders, who serve as subjects of AI, and actors within the value chain (red).
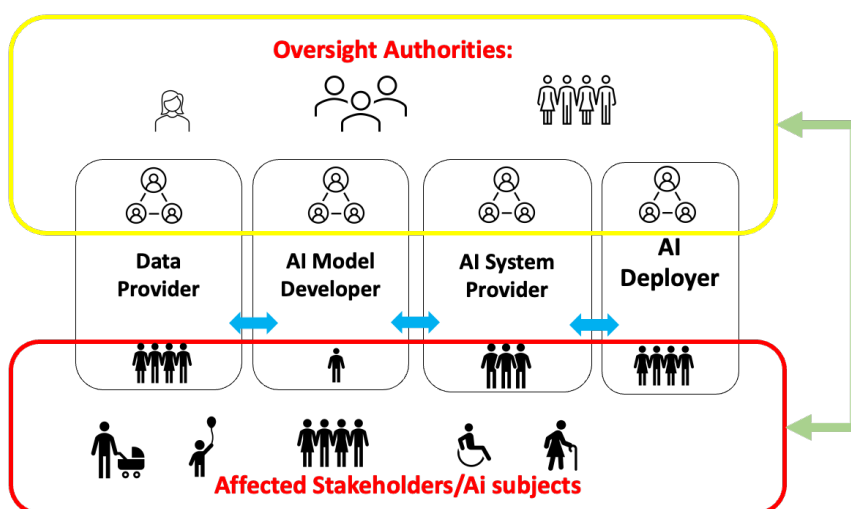● Communication between (affected) societal stakeholders and regulatory actors (green).

Figure 1: AI Communication Lines

Recommendation for AI Governance

This section addresses necessary steps for AI governance in line with a human right based approach to AI and intended for executives in bodies along the socio-technical value chain and regulating bodies. The objective is to foster an ecosystem of excellence by companies and research so that the single market can profit from the product safety approach put forward by the EU AI Act.

01 Emphasize Strategic Investments in National Competent Authorities

In order to achieve alignment and coherence in interactions between regulatory bodies and actors along the value chain, each Member State must make bold investments in building AI competence within their National Competent Authorities (NCAs). This will enable them to effectively handle regulation and establish regulatory sandboxes, ensuring a well-resourced approach to support EU coherence in the AI domain.

02 Ensure Consistent Implementation across Member States

To promote cohesion and avoid regulatory fragmentation, it is crucial to make efforts towards a coherent approach in designating national authorities with competence under the EU AI Act. Harmonization of approaches across Member States is essential to foster EU-wide alignment and prevent disparities in regulatory practices.

03 Empower the AI Office for Effective Enabling

The proposed AI Office must be empowered as a facilitator rather than a mere bureaucracy. Clarity is required regarding its roles within the EU ecosystem. By leveraging emerging bodies from the Digital Services Act (DSA), we can explore ways to enhance the AI Office's competencies and enable it to effectively support AI development and enforce regulatory standards within the European Union.

04 Address the Gaps between the AI Act and IP Law

The intersection between the AI Act and Intellectual Property (IP) law, particularly concerning Generative AI, needs to be addressed comprehensively. Focus areas should include handling training data, disclosure during inference/deployment, and the protection of AI-generated content. Bridging these gaps will provide a more robust regulatory framework for AI innovations.

05 Accelerate the Guidance for Regulatory Sandboxes

To ensure effective trialing and market entry under the EU AI Act, there is a need for detailed provisions regarding regulatory sandboxes. The EU should expedite the issuance of guidance, providing clarity to both national authorities and providers on sandbox implementation and its role in the compliance process.

06 Strengthen Digital Infrastructure for High-Quality Data Processing

To promote AI development, it is crucial to strengthen digital infrastructure for high-quality data processing. This includes involving diverse stakeholders and fostering participation in the infrastructure for testing mechanisms and legal oversight. Additionally, defining "good data"

according to ISO/IEC SC42 5259 (parts 1–5) will ensure data quality for machine learning, aligning with the EU AI Act requirements.

07 Develop Harmonized Standards for Generative AI

Recognizing the multi-purpose nature of Generative AI (GenAI), the EU must develop harmonized standards that address its unique specificities. While most standards are designed for AI systems with specific intended purposes, GenAI necessitates tailored regulations to accommodate its broader applications.

08 Engage Diverse Stakeholders in Foresight for Future-proof Regulation

To create future-proof policy that places a human rights-based approach in the center, it is essential to involve diverse stakeholders in foresight exercises. Active participation from multi-stakeholder fora, including civil society and affected communities, will enable the development of regulations that are inclusive, forward-thinking, and responsive to evolving AI trends and challenges.

Recommendations for building an ecosystem of excellence and trust

The following recommendations address the transformational challenges for actors along the value chain such as Data Provider, AI Model Provider, AI System Provider and AI Deployer. As part of the ecosystem these actors can only succeed and implement developing technologies if the regulatory ecosystem is supporting this aim. This section is designed to gain clarity about roles and potential communication objectives within the AI System to build trust and excellence.

09 Address Unclear Roles and Accountability

The complex nature of value chains in the AI sector has led to unclear accountability and liability, posing risks for human rights abuses. It is crucial to clarify roles within the value chain to ensure a smooth implementation of models from big companies and their regulation.

10 Recognize Active AI Impact on Human Rights

Throughout the AI value chain, fundamental and internationally protected human rights, including human dignity, freedom from discrimination and equality, freedom of expression, right to privacy, and right to life, must be upheld, protected, and respected. AI systems can actively impact human rights, not just passively. Addressing these impacts at every stage of technological development requires commitment from top management and regulators and funding to develop clear guidelines as new impacts emerge.

11 Strengthen Risk Management and Reporting Mechanisms

Corporate stakeholder risk management should go beyond safety concerns and include human rights compliance. Concrete steps are needed to establish robust and ongoing risk management and reporting mechanisms to prevent human rights violations for deployers and providers.

12 Emphasize Transparency and Civil Society Participation

Corporate digital responsibility should involve active engagement with civil society about the ongoing impact of AI systems. Promoting transparency and involving civil society in decision-making processes will foster accountability and public trust.

13 Develop technical and organizational frameworks for Human Right Assessments

In order to effectively address the potential violations of human rights, it is necessary to create impact and risk assessment frameworks that are both technically and organisationally implementable but are also accessible for civil societal and other affected stakeholders. .

14 Integrate Human Rights Assessment Along the Value Chain

Human rights assessment should be a continuous process, integrated into all stages of the AI technical value chain. Centering human dignity and recognizing potential and real harms, such as bias and discrimination, is paramount for building an ecosystem of trust. AI value chain actors should commit to engage with suppliers and customers to recognise and mitigate such harms in a manner that is transparent to the potential subjects of such harm.

15 Ensure Procedural Rights and Accountability

Procedural rights, such as access to redress, meaningful appeal mechanisms, fair and impartial decision-making, and transparency, should be integrated into the AI governance framework.

16 Operationalize Risk Assessment and Human Rights Considerations

Clear guidelines should be established for when ongoing risk assessment, human rights impact assessment, and data protection impact assessment should take place along the technical value chain.

Recommendations for Inclusion and Engaging with Civil Society

For the first time rules on product safety and human rights protection have been combined. The objective is to address the social impact of AI products on the social structures we grow, rely upon, and make our community a livable environment for all. This section therefore suggests pathways of engaging with civil society that make the process of AI development safer for everyone to use.

17 Safeguard Human Rights

Ensuring the protection and respect of human rights for all should be a fundamental pillar in the AI ecosystem. The EU AI Act should place a strong emphasis on upholding human rights throughout the AI value chain, fostering an ethical and responsible AI environment.

18 Enhance Transparency of AI-Generated Content

To empower users and enable informed decision-making, a reliable "AI-generation detection mechanism" must be implemented. Users should be able to discern whether content has been generated by AI or humans, aiding them in navigating diverse sources of information effectively and avoiding information overload and misinformation.

19 Implement Authorship Transparency for Human-Machine Collaborative Content

The EU AI Act should address the transparency of authorship when humans and machines collaborate to produce content, such as documents, papers, or code. Clear guidelines and standards should be established to disclose the contribution of both human and machine elements in collaborative works. Appropriate legal solutions should be developed to sensibly handle various cases of shared authorship between human and machine in practical ways.

20 Identify and Address No-Go Areas

It is essential to have conversations and establish boundaries for AI application, particularly in sensitive and potentially harmful domains like welfare, criminal justice, and border protection, among others. The EU AI Act establishes the principle of prohibited uses of AI, but its implementation should facilitate discussions and agility in defining new no-go areas to prevent potential misuse, as well as unethical and human rights-infringing use of AI technologies.

21 Promote Bottom-Up Participation in Data Governance and Standards

To foster trust and inclusivity, the EU AI Act should encourage bottom-up and grounded participation in data governance and standards development. Engaging diverse stakeholders particularly in the domain of industry standard development will lead to more comprehensive and effective regulations. Human Rights experts, civil society actors, and affected communities should be engaged throughout the process and not simply as an afterthought

22 Collect Stakeholder Feedback for Continuous Improvement

Stakeholder feedback is invaluable for refining AI regulations and adapting to emerging challenges. The EU AI Act should incorporate mechanisms for gathering feedback from various actors in the value chain and society, ensuring continuous improvement of AI governance.

Epilogue

These recommendations aim to strengthen the provisions in the EU AI Act and create a harmonious and responsible AI ecosystem that safeguards human rights, fosters transparency, and actively engages stakeholders in the decision-making process. By adopting these measures, the EU can lead the way in promoting ethical and rights-respecting AI practices and create an ecosystem of trust and excellence.

Contributors

Deirdre Ahern (Trinity College Dublin, IE)

Ivan Bratko (University of Ljubljana, SL)

Sneha Das (Technical University Denmark, DK)

Sylvie Delacroix (Birmingham University, UK)

Lena Kästner (University of Bayreuth,DE)

Till Klein (appliedAI Institute for Europe gGmbh, DE)

David Lewis (Trinity College Dublin, IE)

Anjali Mazumder (Alan Turing Institute, UK)

Petra Molnar (Harvard University and York University, USA & CAN)

Lena Nieper (University of Bayreuth, DE, conference coordinator)

Aldo Faisal (University of Bayreuth & Imperial College London, DE & UK, conference organizer)