

Imagining the AI Landscape after the AI Act

Desara Dushi^{1,†}, Francesca Naretto^{2,3,†} and Francesca Pratesi^{4,*,†}

¹Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

²Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy

³University of Pisa, Largo Pontecorvo 3, 56127, Italy

⁴Institute of Information Science and Technologies - National Research Council, via G. Moruzzi 1, 56124 Pisa, Italy

Abstract

We provide a summary of the second Workshop on Imagining the AI Landscape after the AI Act (IAIL 2023), co-located with the 2nd International Conference on Hybrid Human-Artificial Intelligence (HHAI 2023), held on June 27, 2023 in Munich, Germany.

Keywords

IAIL, IA Act, EU Regulation, Ethics

1. Introduction

In April 2021, the EU Parliament published a proposal on regulating AI systems and services, the AI Act (AIA). After long debates, which led to several amendments to the first draft, the AIA is heading towards the final stages of its approval, after a positive vote by the European Parliament on 11 May 2023, and successfully passing the plenary with an overwhelming majority on 14 June 2023, moving now towards institutional negotiations.

The AI act is a promising initiative that will soon become the world's first comprehensive law on AI. It follows a risk based approach, regulating AI technologies based on their capacity to cause harm, and therefore also banning AI applications that pose an unacceptable risk and imposing a strict regime for high-risk use cases.

The latest developments in technology, such as generative AI, like chatGPT, led to several changes to the text of the AIA before its vote in plenary in June. These changes include the definition of AI, which was closer aligned with the definition of the OECD, the list of prohibited practices was extended to subliminal techniques, biometric categorisation, predictive policing, internet-scraped facial recognition databases, and emotion recognition software banned in law enforcement, border management, workplace and education. Recommender systems of prominent social media were added as high-risk. The list of high-risk areas and use cases,

HHAI-WS 2023: Workshops at the Second International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 26–27, 2023, Munich, Germany


*Corresponding author.


†These authors contributed equally.

✉ francesca.pratesi@isti.cnr.it (F. Pratesi)

🌐 <https://lsts.research.vub.be/en/desara-dushi> (D. Dushi); <https://datasciencephd.eu/students/francesca-naretto> (F. Naretto); <https://kdd.isti.cnr.it/people/pratesi-francesca> (F. Pratesi)

🆔 0000-0002-4260-4473 (F. Pratesi)

 © 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

as well as the obligations for providers of high-risk AI applications were made more precise. A fundamental rights impact assessment and obligation to monitor environmental impact were added to the list of requirements. Ultimately, an AI office was established to coordinate cross-border cases.

At the time of writing the AIA is at the trilogue stage, interinstitutional negotiations between the MEPs, the EU Council of Ministers, and the European Commission. The AIA is planned to be adopted before the end of 2023.

Despite being an EU legislation, therefore applicable only to Member States, the AIA seems to have a clear extraterritorial scope, in that it applies to any AI system or service that has an impact on European Citizens, regardless of where its provider or user is located. Indeed, the AIA, once adopted, will undoubtedly have a significant impact in the EU and beyond.

Almost in parallel with the EU initiative for regulating AI, the Council of Europe (CoE) has been negotiating a draft Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law. Differently from the EU however, the CoE did not make the text of the draft convention public until very recently. However, the EU has been actively participating in the discussions aiming for consistency in the terminology and principles used within the convention¹. Aiming to provide the participants with an overview of AI regulation initiatives beyond the EU, and for being able to make a comparison between the two initiatives, as part of the workshop agenda, we invited for a 'fireside chat' the former chair of the CoE Ad-Hoc Committee on AI (CAHAI) and currently vice-chair of the CoE Committee on AI (CAI).

The purpose of IAIL 2023 was to investigate how will the proposed regulation affect the technological development in EU and how will it impact non-EU tech companies operating in the EU, how to operationalize the ethical requirements of the AI Act, to what extent does the AI act protect individual rights, do we need to define new metrics for validating the goodness of an AI system in terms of privacy, fairness, explainability, and much more.

Topics of interest include, but are not limited to:

- The AI Act and future technologies
- Applications of AI in the legal domain
- Ethical and legal issues of AI technology and its application
- Dataset quality evaluation
- AI and human oversight
- AI and human autonomy
- Accountability and Liability of AI
- Algorithmic bias, discrimination, and inequality
- Trust in practical applications of and data-driven decision-making in AI systems
- Transparent AI
- AI and human rights
- The impact of AI and automatic decision-making on rule of law
- Explainable by design
- Privacy by design

¹See the EU delegation to the Council of Europe: https://www.eeas.europa.eu/delegations/council-europe/council-europe-committee-ai-completes-first-reading-draft-ai-convention_en?s=51

- Fairness by design
- AI risk assessment
- Explainability metrics and evaluation

Papers intended to foster discussion and exchange of ideas. Submissions with an interdisciplinary orientation were particularly welcome, e.g. works at the boundary between machine learning, AI, human-computer interaction, law, digital philosopher, and ethics.

2. Organization

2.1. Workshop Chairs

- Desara Dushi, Vrije Universiteit Brussel (Belgium)
- Francesca Naretto, Scuola Normale Superiore (Italy) and Computer Science Department - University of Pisa (Italy)
- Francesca Pratesi, Institute of Information Science and Technologies - National Research Council (Italy)

2.2. Program Committee

- Costanza Alfieri - University of L'Aquila
- Denise Amram - Scuola Superiore Sant'Anna
- Nertil Bërdufi - University College Beder
- Marco Braghieri - King's College London
- Valeria Caforio - Università Bocconi
- Federica Casarosa - Scuola Superiore Sant'Anna
- Gizem Gezici - Scuola Normale Superiore
- Rūta Liepiņa - Maastricht University
- Cecilia Panigutti - Joint Research Centre
- Giorgia Pozzi - TUDelft
- Clara Punzi - Scuola Normale Superiore
- Giulia Schneider - Università Cattolica del Sacro Cuore
- Mattia Setsu - University of Pisa
- Francesco Spinnato - Scuola Normale Superiore

3. Summary of the workshop

The workshop was highly interdisciplinary and brought together researchers from different backgrounds. The workshop consisted of one keynote speech from Josep Domingo-Ferrer, Distinguished Full Professor of Computer Science and an ICREA-Acadèmia Researcher at Universitat Rovira i Virgili, Tarragona, Catalonia, one fireside chat with Gregor Stojin, Vice Chair of the Committee on Artificial Intelligence at the Council of Europe (CAI, since 2022) and the former Chair of the Ad Hoc Committee on AI (CAHAI, 2019 – 2021), and two sessions of paper presentations with a QA.

3.1. Submissions

The Program Committee (PC) received a total of 7 submissions. Each paper was peer-reviewed by at least three PC members, by following a double-blind reviewing process. The committee decided to accept 6 papers: 2 regular papers (i.e., 12+ pages), 3 short papers, and 1 abstracts. The abstract can contain preliminary or already published work, while papers must contain original work.

3.2. Detailed Program

The IAIL 2023 program was organized in welcome and final remarks sessions, two invited talks, of which one was a classic presentation and the other was a sort of interview, and two paper presentation sessions.

Papers were grouped in two sessions:

Session 1

- Fabienne Ufert and Zachary Goldberg, *How SMEs Ought to Operationalize AI Risk Assessments Under the AI Act*.
- Natália Slosiarová, Matúš Mesarčík, Peter Jurkáček and Juraj Podroužek, *Trustworthy AI in the dental care beyond Artificial Intelligence Act*.
- Karla Aniela Cepeda Zapata, Ritesh Patil, Tomás Ward, Róisín Loughran and Fergal McCaffery, *Analysis of the Classification of Medical Device Software in the AI Act Proposal*.

Session 2

- Maryna Vahabava, *The risks associated with generative AI apps in the European Artificial Intelligence Act (AIA)*.
- Irina Carnat, *Holding the AI Act accountable*.
- Gizem Gezici, Chiara Mannari and Lorenzo Orlandi, *The Ethical Impact Assessment of Selling Life Insurance to Titanic Passengers*.

The papers presentation sessions followed a highly interactive format. They were structured into short presentations with ample room for questions and comments.

3.3. Summary of the presentations

The workshop started with an overview of the goal and the program of the day, and a presentation of the projects which co-founded the event.

Josep Domingo-Ferrer's talk was about ethics-by-design in decentralized settings, relying on Federated Learning, where models are shared instead of data, and he proposed a co-utile approach, where mutual help is the best rational option to take, even for selfish agents, in order to have agents that behave ethically by-design.

In a relaxed interview style (what we call as fireside chat) Gregor Stojin offered a more political point of view, highlighting the different roles of the EU and the Council of Europe

and how this impacts their policy making processes. He also pointed out the needs of multi-stakeholders, to collect inputs from different communities, and evaluate the impact of technology on the population.

In the first paper presented, *How SMEs Ought to Operationalize AI Risk Assessments Under the AI Act*, the focus was on investigating the AI Act from the perspective of Small Medium Enterprises (SMEs), to avoid overburdening SMEs with excessive regulation that could hinder innovation. Indeed, the AI Act also contains some exceptions, intended to foster innovation, for research, testing and development. Then, the presentation provided some practical advice to operationalize the AI Act in SMEs, such as: performing an impact assessment instead of a risk assessment only, i.e., focusing also on benefits and opportunities; scheduling regular meetings with personnel; establishing a clear boundary between the company structure (formal components, processes and procedures) and the content (the ethical risks and benefits aimed to address). The following two papers focused on specific medical case studies, advocating a multi-faceted approach that takes into account not only technical factors, but also ethical principles and human rights considerations together with close engagement of relevant stakeholders. In the first talk, related to the Trustworthy AI in the dental care beyond Artificial Intelligence Act work, authors championed a positive “ethification” of the AI regulation. They also warn on the ethical and social risks of reporting human biases in AI systems, of over-reliance in decision support systems, of providing low transparency, and of the management of very sensitive data. In the second talk, named *Analysis of the Classification of Medical Device Software in the AI Act Proposal*, authors focused on the similarities and the differences between Software as a Medical Device (SaMD), i.e., a software for medical purposes which may not necessarily be part of the hardware, and Software in a Medical Device (SiMD), a software usually embedded into an medical device to support its performance, which could also be utilized as an accessory and combined with other articles.

The second papers session started with *The risks associated with generative AI apps in the European Artificial Intelligence Act (AIA)*, where the author investigated the risk of the Generative Artificial Intelligence (GPAI), like ChatGPT, such as the incorrect use of user data collected, the veracity of the information provided, and the increasing peril of misinformation due to manipulation and distortion of data. The talk continues with the description of the Italian National Data Protection Authority case, which recently asked OpenAI company to provide some clarification on the processing of personal data due to suspected violations of the European Privacy Regulation (GDPR). This was an interesting precedent for the whole European Union; indeed, many other EU countries followed this complaint and moved formal requests to the company, which after a while revised its terms and conditions. The fourth paper to be presented was *Holding the AI Act accountable and it is the only unpublished extended abstract of the workshop*. Here, the author focused on the subtle difference between responsibility, as the duty to carry out the related tasks, accountability, i.e., the fact that the person who undertakes the task is able to give an account, reason or explanation for the action, and liability, as a form of legal or legislative accountability. Finally, with *The Ethical Impact Assessment of Selling Life Insurance to Titanic Passengers*, authors investigated in a concrete (even if fictitious) case study the various ethical requirements, performing a complete ethical impact assessment

and showing the results of the analysis in terms of effectiveness (i.e., true/false positive and true/false negative) of the model, fairness and solidarity, prevention of harms, and explicability. Finally, there was a discussion on the tensions among different ethical dimensions, and the trade-off among them, giving further evidence that there is no one-size-fits-all solution.

4. Conclusion and Remarks

From the discussion carried out in the IAIL 2023 workshop, it appears evident that multidisciplinary is a key point for the effectiveness of the EU legal and ethical framework. The workshop itself is a small evidence of the productive results arising from the dialogue of scholars in the different disciplines, having different approaches and motivation. Engaging in conversations and collaborations on human rights is the main goal that needs to be pursued, in Europe and hopefully even beyond. Other aspects that came up are the importance of taking particular care of generative AI, the problem of many hands in dealing with the accountability principle, and the needs of concrete steps to operationalize the AIA. Finally, some skepticism emerged regarding the exception in the AI-Act for military, defense and national security. However, the papers highlighted the importance and the strength of having an uniform EU legal and ethical framework.

Acknowledgments

This workshop was partially supported by the following:

- SoBigData++ (GA n. 871042) - “European Integrated Infrastructure for Social Mining and Big Data Analytics” (<https://plusplus.sobigdata.eu/>).
- The TAILOR network (GA n. 952215) - Foundations of Trustworthy AI – Integrating Reasoning, Learning and Optimization (<https://tailor-network.eu>).
- HumanE-AI-Net (GA n. 952026) - European network of Human-Centered Artificial Intelligence (<https://www.humane-ai.eu/>).
- COHUBICOL (GA. n. 788734) “Counting as a Human Being in the Era of Computational Law” (<https://www.cohubicol.com/>).

Authors would like to thank the HHAI 2023 workshop chairs and organization for providing an excellent framework for IAIL 2023.