Fourth European Workshop on Algorithmic Fairness (EWAF'25)

HILDE WEERTS, Eindhoven University of Technology, The Netherlands
MYKOLA PECHENIZKIY, Eindhoven University of Technology, The Netherlands
DORIS ALLHUTTER, Austrian Academy of Sciences, Austria
ANA MARIA CORRÊA, KU Leuven, Belgium
THOMAS GROTE, University of Tübingen, Germany
CYNTHIA LIEM, Delft University of Technology, The Netherlands

The European Workshop on Algorithmic Fairness (EWAF) aims to foster dialogue between researchers working on algorithmic fairness in the context of Europe's legal and societal framework, especially in light of the European Union's attempts to promote ethical AI and the turn to AI for the common good. EWAF welcomes submissions from multiple disciplines, including but not limited to computer science, law, philosophy, and social science, as well as interdisciplinary and transdisciplinary work. The fourth edition (EWAF'25) will provide a dedicated venue for interdisciplinary discourse on algorithmic fairness via keynotes, paper presentations, a panel discussion, and interactive sessions.

Reference Format:
Hilde Weerts, Mykola Pechenizkiy, Doris Allhutter, Ana Maria Corrêa, Thomas Grote, and Cynthia Liem. 2025. Fourth European Workshop on Algorithmic Fairness (EWAF'25). In *Proceedings of Fourth European Workshop on Algorithmic Fairness (EWAF'25).* Proceedings of Machine Learning Research, 9 pages.

1 Introduction

The Fourth European Workshop on Algorithmic Fairness (EWAF'25) aims to foster dialogue between researchers working on algorithmic fairness in the context of Europe's legal and societal framework, especially in light of the European Union's attempts to promote ethical AI and the turn to AI for the common good. Building upon the success of the first three editions, EWAF'25 will be held from June 30 to July 2, 2025, at Eindhoven University of Technology (the Netherlands).

2 Objectives and Areas of Interest

The goals of EWAF'25 are to provide a space for algorithmic fairness research grounded in the European perspective and setting, fostering interdisciplinary scholarship and dialogue, and to provide a stimulating scientific program.

Authors' Contact Information: Hilde Weerts, Eindhoven University of Technology, Eindhoven, The Netherlands, h.j.p.weerts@tue.nl; Mykola Pechenizkiy, Eindhoven University of Technology, Eindhoven, The Netherlands, m.pechenizkiy@tue.nl; Doris Allhutter, Austrian Academy of Sciences, Vienna, Austria, doris.allhutter@oeaw.ac.at; Ana Maria Corrêa, KU Leuven, Leuven, Belgium, anamaria.correaharcus@kuleuven.be; Thomas Grote, University of Tübingen, Tübingen, Germany, thomas.grote@uni-tuebingen.de; Cynthia Liem, Delft University of Technology, Delft, The Netherlands, c.c.s.liem@tudelft.nl.

EWAF'25, June 30–July 02, 2025, Eindhoven, NL

EWAF welcomes submissions from multiple disciplines, including but not limited to computer science, law, philosophy, and social science. The workshop also explicitly welcomes submissions of interdisciplinary and transdisciplinary work, sourcing from multiple disciplinary areas and/or highlighting joint insight-building with relevant non-academic stakeholders. A non-exhaustive list of themes includes:

- Industry experiences in developing and implementing fairness interventions, developing standards and practical approaches to introducing fairness in digital innovation governance.
- Empirical and theoretical perspectives from social sciences on fairness and discrimination in Europe (e.g., analysis of labor markets, the concepts of class, race, disability, and discrimination against minorities in different social contexts, intersectional inequality).
- Case studies based on concrete European instances of algorithmic design and regulation that machine learning scholars or practitioners have encountered in their work (e.g., datasets or audits of automated decision-making systems that are used in Europe).
- An analysis of the implications of the European legislative framework for the debate on fairness in machine learning and AI more broadly (e.g., specificities connected to anti-discrimination and data collection legislation and the emerging regulatory frameworks for platforms and AI).
- Principled arguments for certain fairness concepts and measures in specific contexts.
- Implementing fairness in deployed systems, selecting fairness definitions and designing auditing processes.
- Explorations of the relationship and trade-offs between fairness and transparency in practice.
- Fairness and transparency of black-box models.
- Generative AI and fairness, esp. relating to the job market and the data supply chain.
- Studies on the systemic risks of large online platforms (such as social media platforms and online marketplaces) and search engines, in particular regarding protection of minors, dissemination of illegal products, gender-based violence, freedom of expression and disinformation.

3 Program

3.1 Accepted Papers

EWAF'25 invited both full papers and extended abstracts. This year, we received a record number of 96 submissions, of which 66 extended abstracts and 30 full papers. In total, 69 papers were accepted, of which 54 extended abstracts and 15 full papers. All accepted papers will be presented during either a lightning round talk, poster session, or in-depth presentation.

3.2 Interactive Sessions

Recognizing the crucial role of diverse perspectives in discussions on algorithmic fairness, EWAF seeks contributions that extend beyond the academic discourse, involving voices from industry, civil society, government, NGOs, journalists, artists and other non-academic sectors. To this end, we invited interactive sessions (e.g. workshops, panels, unconferences, moderated dialogues, discussion groups, fish bowls, artistic interventions) to address gaps between theory and practice, research and policy, and multiple disciplines/fields tackling ethical AI; position European institutions, both research and governmental, and their impact on ethical AI; foster conversations between

the different stakeholders involved in ethical AI, promoting interdisciplinary and cross-practice discussions; and provide space for discussion and reflection within the algorithmic fairness research community.

This year's edition includes the following accepted interactive sessions:

- "Redefining AI Fairness Through an Indigenous Lens" by Myra Colis
- "Contribute to the technical specification for responsible use of risk profiling algorithms used by the Dutch government" by Laura Muntjewerf, Willy Tadema
- "Designing a 'fair' human-in-the-loop" by Isabella Banks, Jacqueline Kernahan
- "Beyond the Buzzwords: Co-Creating Accountability through Legal and Technological Perspectives on Algorithmic Transparency" by Maria Lorena Flórez Rojas
- "Fairness for whom? On the need for co-creative pathways in AI policy and research" by Ilina Georgieva, Paul Verhagen, Courtney Ford
- "Practices of resistance against algorithmic risk scoring and automated decision-making in welfare" by Doris Allhutter, Karolina Sztandar-Sztanderska

3.3 Keynotes

3.3.1 "Reassembling the Black Box(es) of Algorithmic Fairness: Methods, Interactions and Politics" by Juliane Jarke

Research on algorithmic fairness promotes a view on algorithmic systems as black boxes that need to be "opened" and "unpacked" to make their inner workings accessible. Understanding the black box as a mode of inquiry and knowledge making practice (rather than a thing), I will explore what exactly researchers and practitioners aim to unpack when they examine algorithmic black boxes, what they consider to be constitutive elements of these black boxes, and what is othered or perceived as "monstrous". Following scholarship in the social studies of science and technology (STS), the notion of the monster captures what is considered irrelevant for the constitution and inner workings of an algorithmic black box. It is what is excluded and escapes analysis. I will argue, however, that attention to the outer limits of algorithmic black boxes allows us to explore how social actors, temporalities, places, imaginaries, practices, and values matter for knowledge making about algorithmic fairness. In my talk, I will review three modes of assembling black boxes of machine learning (ML)-based systems: (1) the black box of ML data, (2) the black box of ML algorithms and trained models and (3) the black box of ML-based systems in practice. In reassembling these three distinct ML black boxes, I demonstrate how generative engagements with algorithmic black boxes and their monsters are for the critical inquiry of algorithmic fairness.

Biography. Juliane Jarke is Professor of Digital Societies at the University of Graz. Her research attends to the transformative power of algorithmic systems in the public sector, education and for ageing populations. It intersects critical data & algorithm studies, participatory (design) research and feminist STS. Juliane received her PhD from Lancaster University and has a background in Computer Science, Philosophy, and STS. She has recently co-edited a special issue on Care-ful Data Studies (Information, Communication and Society). Her latest co-edited books include Algorithmic Regimes: Methods, Interactions and Politics (Amsterdam University Press) and Dialogues in Data Power: Shifting Response-abilities in a Datafied World (Bristol University Press). Juliane is also co-organiser

of the Data Power Conference series and Co-PI in the research unit Communicative AI: The Automation of Societal Communication

3.3.2 "On fairness and validity in AI-assisted hiring" by Julia Stoyanovich

In this talk, I will examine the widespread use of AI systems in hiring and employment, highlighting where these tools can be helpful and where they raise concerns around discrimination and validity. I will focus on three lines of work.

First, I will provide an overview of fairness in ranking, offering a perspective that connects formal definitions and algorithmic techniques to the value frameworks that motivate fairness interventions, and to the technical choices that shape the behavior and outcomes of these methods. Second, I will discuss algorithmic recourse, which aims to help individuals reverse negative decisions—such as being screened out by an automated hiring system. I will highlight recent work exploring how resource constraints (i.e., a limited number of favorable outcomes) and competition influence the reliability and fairness of recourse over time. Third, I will present findings from an audit of two commercial systems — Humantic AI and Crystal — which claim to infer job-seeker personality traits from resumes and social media data. I will describe our audit methodology and show that both systems exhibit instability in key measurement facets, rendering them unsuitable as valid instruments for pre-hire assessment.

I will conclude with a discussion of emerging legal and regulatory developments in the U.S. aimed at curbing the unaccountable use of AI in hiring, and reflect on what it would take to ensure these systems are safe, fair, and socially sustainable in this critical domain.

Biography. Julia is an Institute Associate Professor of Computer Science and Engineering at the Tandon School of Engineering, Associate Professor of Data Science at the Center for Data Science, and Director of the Center for Responsible AI. Julia's goal is to make "Responsible AI" synonymous with "AI". She works towards this goal by engaging in academic research, education and technology policy, and by speaking about the benefits and harms of AI to practitioners and members of the public.

Julia's research interests include AI ethics and legal compliance, and data management and AI systems. In addition to academic publications, she has written for the New York Times, the Wall Street Journal, and Le Monde. Julia has been teaching courses on responsible data science and AI to students, practitioners and the general public. She is a co-author of "Data, Responsibly", an award-winning comic book series for data science enthusiasts, and "We are AI", a comic book series for the general audience.

Julia is engaged in technology policy and regulation in the US and internationally, having served on the New York City Automated Decision Systems Task Force, by mayoral appointment, among other roles.

Julia received her M.S. and Ph.D. degrees in Computer Science from Columbia University, and a B.S. in Computer Science and in Mathematics & Statistics from the University of Massachusetts at Amherst. She is a recipient of the NSF CAREER Award and a Senior Member of the ACM.

3.3.3 "Algorithmic Governance in Europe: Inside the Digital Machinery of Discrimination" by Raphaële Xenidis

This talk examines the rise of the digital welfare state in Europe and its implications for social justice, focusing particularly on the discriminatory mechanisms of algorithmic governance. While presented as tools for modernizing public administration and optimizing resource allocation, automated decision-making systems used in welfare programs often institutionalize exclusion, amplify existing inequalities, and erode fundamental rights. Drawing

on cases from Europe, the talk highlights how semi-automated fraud detection systems operate opaquely and ineffectively, disproportionately targeting marginalized populations under the guise of efficiency and fraud prevention. Central to the argument is the application of an intersectional lens — rooted in Black feminist thought — which reveals how these systems perpetuate complex forms of discrimination along axes such as gender, race, disability, and economic status. The analysis critiques existing legal frameworks for failing to address systemic and intersectional forms of algorithmic discrimination and calls for a rethinking of legal interventions.

Biography. Raphaële Xenidis is assistant professor in European law at Sciences Po Law School in Paris, France. Her current research focuses on algorithmic bias and discrimination in automated decision-making systems in the context of European equality law. Raphaële holds a PhD in law from the European University Institute and is an Honorary Fellow at the University of Edinburgh Law School and a Global Fellow at iCourts, University of Copenhagen.

3.3.4 "Algorithmic Fairness, Intersectionality, and Uncertainty" by Johannes Himmelreich

In this talk I examine a fundamental dilemma facing intersectional algorithmic fairness in practice. As regulatory frameworks like the EU AI Act require intersectional approaches to fairness, we confront two interrelated problems that threaten meaningful implementation.

First, the problem of statistical uncertainty: When considering intersectional groups—as opposed to groups defined by a single (demographic) attribute—the number of groups increases exponentially while sample sizes decrease dramatically. This creates statistical uncertainty that renders standard fairness metrics meaningless at best, or morally problematic at worst. Second, the problem of ontological uncertainty: The question of which groups warrant fairness consideration remains theoretically underdetermined. The needed ontological theory would yield G, the set of all groups that are to be included in a fairness audit. The options for such theories include that G consist of (a) groups generated from all possible combinations of protected attributes, (b) some relevant combination, or (c) some relevant groups, such as those groups with a history of disadvantage. This theoretical ambiguity enables "fairness gerrymandering", that is, strategically defining G to achieve desired outcomes.

These problems generate a dilemma. If we include all intersectional groups, statistical uncertainty makes meaningful fairness audits impossible. If we restrict attention to select groups, we face a tension between accommodating ontological uncertainty on the one hand and preventing fairness gerrymandering on the other. Rather than viewing this as grounds for abandoning intersectional fairness, I hypothesize that more work is needed to identify: (1) new approaches to fairness auditing that explicitly account for statistical uncertainty about small groups, and (2) clearer theoretical principles for determining relevant group ontologies.

Biography. Johannes Himmelreich is a philosopher who teaches and works in a policy school. He is an Assistant Professor in Public Administration and International Affairs in the Maxwell School at Syracuse University. He works in the areas of political philosophy, applied ethics, and philosophy of science. Currently, he researches the ethical quandaries that data scientists face, how the government should use AI, and how to check for algorithmic fairness under uncertainty.

He published papers on "Responsibility for Killer Robots," the trolley problem and the ethics of self-driving cars, as well as on the role of embodiment in virtual reality.

He holds a PhD in Philosophy from the London School of Economics (LSE). Prior to joining Syracuse, he was a post-doctoral fellow at Humboldt University in Berlin and at in the McCoy Family Center for Ethics in Society at Stanford University. During his time in Silicon Valley, he consulted on tech ethics for Fortune 500 companies, and taught ethics at Apple.

3.4 Panel

3.4.1 "Algorithmic citizen profiling and risk-scoring in the welfare state: civil society perspectives and litigation", moderated by Doris Allhutter & Karolina Sztandar-Sztanderska

The introduction of algorithmic citizen profiling and risk-scoring for purposes such as welfare fraud detection and the allocation of social services has become a widespread practice across Europe. While governments, for instance, introduce laws that oblige social security bodies to prevent social fraud, civil society organizations, data journalists and researchers raise concerns that biased, non-transparent automated risk assessment and profiling affect citizens' privacy and social rights. NGOs demand insight into algorithmic practices used to segment access to welfare or to flag citizens in vulnerable life situations as potential fraudsters. In the Netherlands and France, NGOs and coalitions between data rights organizations, organizations representing disadvantaged groups and unions have taken legal action to defend equal access to social security and the right of citizens to social protection.

This panel brings together representatives from different organizations, interest groups and journalism to discuss how algorithmic systems affect data protection and social rights and to deliberate on strategies that have been successful in gaining transparency and organizing resistance. After the panel discussion, the audience is invited to join the exchange on cases and experiences from different national contexts. The aim of this interactive session is to facilitate a dialogue between the algorithmic fairness community, civil society groups and data journalists to fathom organizing within and across national contexts.

Panelists.

- Bastien Le Querrec (La Quadrature du Net)
- Soizic Penicaud (Observatory of Public Algorithms - Odap.fr)
- Tijmen Wisman (Stichting Platform Bescherming Burgerrechten, Vrije Universiteit Amsterdam)
- Mateusz Wrotny (Panoptykon Foundation)

4 Organization

EWAF'25 was organized by the following people.

General Chairs

- Hilde Weerts, Eindhoven University of Technology
- Mykola Pechenizkiy, Eindhoven University of Technology

Program Chairs

- Doris Allhutter, Austrian Academy of Sciences
- Ana Maria Corrêa, KU Leuven
- Thomas Grote, University of Tübingen

- Cynthia Liem, Delft University of Technology

Interactive Sessions Chair

- Linnet Taylor, Tilburg University

Proceedings Chairs

- Mattia Cerrato, Johannes Gutenberg-Universität Mainz
- Marco Favier, University of Antwerp

Local Organizing Committee

- Gizem Cenç, Eindhoven University of Technology

Area Chairs

- Agathe Balayn, Microsoft Research
- Alessandro Fabris, University of Trieste
- Jose M. Alvarez, KU Leuven
- Maarten Buyl, Universiteit Ghent
- Sune Holm, Copenhagen University

Program Committee

- Alesia Vallenas Coronel, Johannes-Gutenberg Universität Mainz
- Ana-Andreea Stoica, Max Planck Institute for Intelligent Systems
- Antonio Mastropietro, University of Pisa
- Antonio Vetrò, Polytechnic Institute of Turin
- Ayan Majumdar, MPI-SWS
- Blanca Luque Capellas, Johannes-Gutenberg Universität Mainz
- Carlotta Rigotti, Leiden University
- Chadha Degachi, Delft University of Technology
- Charlotte Ducuing, Katholieke Universiteit Leuven
- Chiara Ullstein, Technische Universität München
- Christoph Kern, University of Munich, Ludwig-Maximilians-Universität München
- Clara Rus, University of Amsterdam, University of Amsterdam
- Daphne Lenders, School of Education Pisa
- David Hartmann, Technische Universität Berlin
- Deborah D. Kanubala, Universität des Saarlandes
- Eike Petersen, Fraunhofer Institute for Digital Medicine MEVIS
- Emmanouil Krasanakis, CERTH/ITI
- Fabian Fischer, Austrian Academy of Sciences
- Fariba Karimi, Technische Universität Graz
- Fatemeh Mohammadi, University of Milan
- Francesco Nappo, Polytechnic Institute of Milan

- Francien Dechesne, Leiden University
- Frederic Gerdon, Universität Mannheim
- Giandomenico Cornacchia, International Business Machines
- Guillaume Bied, Universiteit Gent
- Iris Dominguez-Catena, Universidad Pública de Navarra
- Irmak Erdogan-Peter, KU Leuven
- Jan Grenzebach, Federal Institute for Occupational Health and Safety
- Jan Simson, Ludwig-Maximilians-Universität München
- Jiaxu Zhao, Eindhoven University of Technology
- Karima Makhlouf, University of Doha for Science and Technology
- Kristen Marie Scott, KU Leuven
- Laura State, University of Pisa
- Laurens Naudts, University of Amsterdam
- Ludwig Bothmann, Ludwig-Maximilians-Universität München
- Marco Favier, Universiteit Antwerpen
- Marco Rondina, Polytechnic Institute of Turin
- Marcos L P Bueno, Radboud University
- Marina Ceccon, University of Padua
- Marius Köppel, ETH Zurich
- Marta Marchiori Manerba, University of Pisa
- Martina Cinquini, University of Pisa
- MaryBeth Defrance, Universiteit Gent
- Michele Loi, Algorithmwatch
- Miriam Fahimi, Alpen-Adria Universität Klagenfurt
- Nathan Genicot, Université Libre de Bruxelles
- Nikolaus Poechhacker, Karl-Franzens-Universität Graz
- Peter Müllner, Technische Universität Graz
- Raphael Romero, Universiteit Gent
- Rianne M. Schouten, Eindhoven University of Technology
- Roger Campdepadros, University of Girona
- Sandro Radovanović, University of Belgrade
- Sebastian Zezulka, Eberhard-Karls-Universität Tübingen
- Shelly Yiran Shi, University of California, San Diego
- Shiying Li, University of Wisconsin - Madison
- Simon Egbert, University of Massachusetts at Amherst
- Simon Schaupp, Technische Universität Berlin
- Simone Casiraghi, Vrije Universiteit Brussel
- Sofie Goethals, Columbia University
- Solal Nathan, Université Paris-Saclay
- Teresa Scantamburlo, University of Venice

Acknowledgments

EWAF'25 was made possible by the generous support of the Dutch Research Council (NWO Scientific Meetings and Consultations Grant), the European Centre for Algorithmic Transparency (ECAT), the Dutch Data Protection Authority (AP), the Netherlands Research School for Information and Knowledge Systems (SIKS), the Eindhoven Artificial Intelligence Systems Institute (EAISI), and the TU/e Center for Safe Artificial Intelligence (SafeAI).