

ECAT Research Workshop AGENDA

27 November 2024

Walter Hallstein Room, Berlaymont

The European Centre for Algorithmic Transparency was established in 2023 as a hub of technical and scientific expertise for the enforcement of the Digital Services Act (DSA). Since launching, our team of dedicated researchers has been working to shine a light on the long-running impact of algorithms and uncover risks represented by online platforms, alongside their potential mitigations.

These insights inform policy, strengthen investigations under the DSA and contribute to the public discussion.

Exchanging ideas and staying up to date with the work of researchers is core to our role as a centre of expertise. By pooling knowledge, we improve our ability to understand risks associated with platforms ubiquitous in the lives of EU citizens, putting us in a better position to help mitigate such risks.

At this workshop, we will listen and learn from those carrying out cutting-edge research on the online world and its effects, and share lessons from our own work so far.

9:00-09:30 **Registration and coffee**

09:30-09:45 **Welcome**

Alberto Pena Fernandez, Head of Unit, European Centre for Algorithmic Transparency

09:45-11:15 **Algorithmic auditing: Can it foster meaningful accountability and progress?**

The role of independent algorithmic auditing for regulatory compliance is set to grow significantly in coming years, and with it, a budding new sector. In this session, experts will share their perspectives on the most robust methods for algorithmic auditing, and how a high quality can be maintained across providers.

We will address what some of the main challenges are for algorithmic auditors, and what regulators should look out for when trying to ensure that an audit presents reliable results and can be used effectively to hold the audited organisation to account.

The session will also focus on some of the elements audited in the DSA, specifically the concept of systemic risks.

Chairs: **Dr Josep Soler Garrido**, DSA Enforcement Lead, European Centre for Algorithmic Transparency and **Dr Cecilia Panigutti**, Scientific Project Officer, European Centre for Algorithmic Transparency

Speakers

Dr Gilles Tredan, Researcher, CNRS

Dr Gemma Galdon Clavell, CEO, Eticas

Dr Petros Terzis, Postdoctoral Researcher, University of Amsterdam

Dr Oliver Marsh, Head of Tech Research, AlgorithmWatch (in-person) **and** Dr Michele Loi, Senior Science Manager, AlgorithmWatch (online)

11:15-11:35 **Coffee break**

11:35-13:00 **The promise and challenges of automated content moderation**

According to the DSA Transparency Database, half of reported content moderation decisions in the EU are fully automated. How well it works therefore hugely affects what we see on online platforms, as well as how any content we share ourselves is treated.

In this session, we ask what the most common technologies and tools for content moderation are, and what their benefits and limitations could be. We will explore the use of generative AI for moderation purposes, as well

as the balance between human moderators and automated tools in the overall approach of platforms.

Through the discussion, we also look to gain a better understanding of any challenges presented by specific types of content, and if automated approaches are better suited for some than others.

Chair: **Dr Yves Punie**, Deputy Head of Unit, European Centre for Algorithmic Transparency

Speakers

Dr Viviana Patti, Associate Professor of Computer Science, University of Turin

Dr Robert Gorwa, Postdoctoral Researcher, WZB Berlin Social Science Center

Professor Claudio Palazzi, Professor of Computer Science **and** Mirko Franco, PhD student, University of Padua

13:00-14:00

Lunch

14:00-15:30

Being a woman online: The prevalence of online gender-based violence and how to combat it

Online experiences are not the same for everyone. Depending on who you are, you may be at greater risk of unpleasant experiences, harassment or even violence.

In this session, we will hear from experts on online gender-based violence (OGBV), sharing their knowledge of its prevalence and how this has developed over time. We will discuss the role of the manosphere and male supremacy, the spread of deepfake technologies, and the risks online platforms present when it comes to exacerbating OGBV and misogynistic content.

Finally, we explore possible interventions by platforms to prevent and combat OGBV, and what platform data may enable researchers to help come up with such solutions in the future.

Chair: **Dr Eleonora Esposito**, Case Handler Officer, DG Connect

Speakers

Professor Clare McGlynn, Professor of Law, Durham University

Dr Silvia Semenzin, Postdoctoral Fellow in Digital Sociology, University Complutense of Madrid

Dr Lisa Sugiura, Associate Professor, Portsmouth University

15:30-15:50 **Coffee break**

15:50-17:20 **Recommender systems for good and bad? What we know about platform algorithms and their impact**

The role of recommender systems in shaping our experience of the online world is hotly debated. For this last session of the workshop, we will dive into what we actually know about the impact of these algorithms, especially when it comes to any proliferation of illegal and potentially harmful content. We explore how robust the evidence is, and what platform design choices may increase or decrease such risks.

We are also going to reflect on how recommender systems should be explained to users, and what types of data would help future studies on their impact.

Chair: **Dr Emilia Gomez**, Lead Scientist, European Centre for Algorithmic Transparency

Speakers

Professor Jana Lasser, Professor for Data Analysis, University of Graz

Dr Lien Michiels, Researcher, University of Antwerp & Vrije Universiteit Brussels

Dr Yashar Deldjoo, Assistant Professor, Polytechnic University of Bari

17:20-17:30 **Close**

Alberto Pena Fernandez, Head of Unit, European Centre for Algorithmic Transparency

Speaker bios

Yashar Deldjoo is a senior research scientist and a tenure-track Assistant Professor at the Polytechnic University of Bari, Italy. He serves as a senior program committee (SPC) for major conferences such as SIGIR, CIKM, ECAI, and WebConf, and is an associate editor for ACM Computing Surveys (CSUR). Yashar earned his Ph.D. with distinction in computer science from Politecnico di Milano (Italy) and an M.Sc. in electrical engineering from Chalmers University of Technology (Sweden). His research focuses on recommender systems, and integrating the main trustworthy AI principles such as fairness, adversarial robustness, and explainability into the recommendation systems. His recent research focuses on exploring the benefits and risks of Generative AI and Large Language Models (LLMs) in recommender systems, with particular emphasis on their societal impact.

Mirko Franco is a PhD student in Brain, Mind, and Computer Science at the Department of Mathematics of the University of Padua. He received his M.Sc. degree in Computer Science in 2021, cum laude, and completed his B.Sc. degree in Computer Science in 2019 at the same university. His main research interests include content moderation, online safety, and mobile systems. Mirko is a member of the Technical Program Committee of the ACM Conference on Information Technology for Social Good (GoodIT), and of the IEEE Consumer Communications and Networking Conference (CCNC). His paper “Characterizing Non-Consensual Intimate Image Abuse on Telegram Groups and Channels” won the Best Paper Award at the 4th International Workshop on Open Challenges in Online Social Networks.

Dr Gemma Galdon-Clavell is a pioneer and global force in AI safety and auditing, ensuring that machine learning tools truly serve society. She is the founder and CEO of Eticas.ai, a venture-backed organization that identifies, measures and corrects algorithmic vulnerabilities, bias and inefficiencies in predictive and LLM tools. Eticas’ software, the ITACA platform, is the first solution to automate impact analysis and monitoring, ensuring that AI systems are high performing and safe, explainable, fair and trustworthy.

Dr. Galdon-Clavell’s impactful work – and passion for disrupting the status quo – earned her recognition as a Mozilla Rise25 honoree in 2024, Hispanic Star Awardee at the United Nations in 2023, an Ashoka Fellow in 2020 and a finalist at the EU Prize for Women Innovators awarded by the European Commission in 2017. In 2023 the BBC acknowledged her as one of the “people changing the world” and in 2024 she was honored by Forbes Women as one of the “35 Leading Spanish Women in Technology”, praised as “a pioneer in algorithmic auditing software”.

Robert Gorwa is a research fellow at the Berlin Social Science Center (WZB). He has published widely on platform governance, content moderation, and other emerging technology policy issues, and his first book, *The Politics of Platform Regulation*, was published in 2024 by Oxford University Press.

Jana Lasser is a Professor for Data Analysis at University of Graz where she leads the research group of Complex Social & Computational Systems at the interdisciplinary center IDEa_Lab. She researches emergent phenomena in complex social systems, employing methods from machine learning, data science, natural language processing and computational and statistical modelling to understand how humans behave in socio-technical environments. Her current research interests include the effectiveness of counterspeech strategies and the spread

of misinformation on social media platforms, the fracturing of our society's understanding of "honesty" and the impact of social media recommendation algorithms on societal outcomes.

Michele Loi is a distinguished technology ethics expert combining rigorous academic research with practical policy impact in AI governance and digital rights. Previously led multiple research projects as Principal Investigator at the University of Zurich's Digital Society Initiative and served as a Marie Skłodowska-Curie Individual Fellow at Politecnico di Milano. Now driving change as Senior Science Manager at AlgorithmWatch, bridging scholarly insights with actionable policy recommendations and corporate consulting on AI ethics. Brings deep interdisciplinary expertise in analyzing societal implications of digital technologies, with demonstrated success in engaging diverse stakeholders across private sector, NGOs, government, and academia. Specializes in translating complex AI ethics research into practical governance frameworks and comprehensive policy solutions.

Oliver Marsh

As Head of Tech Research, Oliver leads AlgorithmWatch's research work and partnerships on policy areas including the Digital Services Act and the AI Act. He is also responsible for integrating the research strategy into campaigning and advocacy. Oliver previously worked on platform and data governance as an official in Downing Street in the UK, and as an analyst of online harms for CASM Technology, The Institute for Strategic Dialogue, Demos, and The Tony Blair Institute. He holds a PhD in sociology of social media from University College London and a degree in Natural Sciences & History and Philosophy of Science from Cambridge University.

Professor Clare McGlynn is a Professor of Law at Durham University, UK, and an expert on violence against women and girls. Her pioneering research developing the concept of '[image-based sexual abuse](#)' has influenced new criminal laws around the world and she regularly advises parliaments, civil society and internet platforms including Meta, Google and Bumble. She worked closely with the Council of Europe and EDVAW to draft its first [Thematic Paper on the Digital Dimension of VAW](#) and she is a member of the Council of Europe's [Expert Committee](#) on Tech-Facilitated Violence Against Women. She [worked with HateAid](#) and Bumble on the EU's Gender-Based Violence Directive and the DSA. Her research on the sexually violent nature of mainstream porn (with Vera-Gray) is the largest [study](#) to date of online porn content. In 2021 she was awarded an honorary PhD from Lund University. She is co-author of [Cyberflashing: recognising harms, reforming laws](#) (2021), [Image-Based Sexual Abuse: a study on the causes and consequences of non-consensual imagery](#) (2021) and author of [Families and the European Union: law, politics and pluralism](#) (2006) and [The Woman Lawyer: making the difference](#) (1998). [@McGlynnClare](#) [www.ClareMcGlynn.com](#) [LinkedIn](#)

Lien Michiels is an interdisciplinary researcher who splits her time between imec-SMIT (Comm. Sci) at Vrije Universiteit Brussel, and the Adrem Data Lab (Comp. Sci.) at the University of Antwerp. She obtained a PhD in Computer Science from the University of Antwerp for her work on evaluation of recommender systems.

In her ongoing work, she continues to explore how we can evaluate the impact of recommender systems, for example, on users' experiences of serendipity or filter bubbles on online platforms. Lien is also a creator and maintainer of RecPack, an open-source package for evaluating recommendation algorithms offline.

Claudio E. Palazzi, PhD in Computer Science, is Associate Professor of Computer Science at the University of Padua (Department of Mathematics). He co-leads the MobileLab Research Group and serves as the Head of the Computer Science Bachelor's and Master's programs.

His research primarily focuses on computer networks, with a special emphasis on mobile applications, multimedia entertainment, social networks, and the application of IT technologies for social good. He is one of the founders of ACM GoodIT (International Conference on Information Technology for Social Good) and serves as Associate Editor for IEEE Transactions on Multimedia and Elsevier Computer Networks. Over the years, he has been the Principal Investigator of numerous research projects funded by the Italian, EU, and U.S. governments.

Viviana Patti, PhD in Computer Science, is Associate Professor of Computer Science at the University of Turin (Computer Science Department) and part of the scientific board and executive committee of the Center for Logic, Language, and Cognition (LLC). She is currently vice-president of the board of directors of the Italian Association of Computational Linguistics (AILC). She has been member of the Guarantee Committee of the University of Turin, which develops and carries out actions to promote equal opportunities, well-being in the workplace and non-discrimination. Her main research interests are in the areas of Natural Language Processing, Computational Linguistics and Affective Computing, and include sentiment analysis, the automatic recognition of emotions and irony, with a particular focus on social media texts and the relationship between language and social structure. Recently, she has been applying her research to the automatic monitoring of online hate speech, with a particular focus on hate speech against migrants, populist rhetoric, and misogyny.

Silvia Semenzin is a research fellow in Digital Sociology at the Complutense University of Madrid and a feminist activist. She works on online misogyny, technology-facilitated gender-based violence and emerging technologies. At the end of 2018 she became the promoter of the #in-timitàviolata campaign that led to the recognition in Italy of a law against the non-consensual sharing of intimate images. She is currently working as a contractor for the EU Commission, focusing on enforcing the Digital Services Act (DSA) on pornographic platforms.

Dr Lisa Sugiura is Associate Professor in Cybercrime and Gender in the School of Criminology and Criminal Justice at the University of Portsmouth and Fellow and Chair of the Board of Directors for the Institute for Research on Male Supremacism. Lisa is an internationally recognised expert in online gender-based violence and technology-facilitated sexual violence and regularly appears on national and international media news outlets discussing these subjects. Her research projects, which include funding from the UK National Cyber Security Centre (NCSC) and the Home Office, involve the language of cybersexism, victims of computer misuse, the role of technology in domestic abuse, and extremist and misogynistic behaviours in manosphere and incel communities. Lisa has published extensively on topics including rape culture, incels (involuntary celibates), and online misogyny. She is the author and co-editor of the books: *The incel rebellion: the rise of the manosphere and the virtual war against women* Emerald publishing, and the *Palgrave Handbook of Gendered Violence and Technology*.

Petros Terzis is a Postdoctoral Researcher at the Institute for Information Law, University of Amsterdam. His work and publications focus on the regulation of computational infrastructures and the role of law in the global value chains of technology. Today, working alongside Prof Joris van Hoboken, he is exploring the political economy of computing with an emphasis on early developments in quantum technologies and their standardisation. From 2021 to 2023, Petros was a Research Fellow at the Faculty of Laws, University College London (UCL), where (along with Dr Michael Veale) he studied the problem of infrastructural programmability from a legal-political perspective. He holds a PhD in Information Law from the University of Winchester (UK) and a

Master's degree in Law and Economics from the University of Macedonia (Greece). As an accredited Fellow of the Higher Education Academy of the UK (FHEA) he has taught a wide range of subjects at the intersection of Law and Technology at UCL, the University of Amsterdam, and the London School of Economics.

Gilles Tredan

Since November 2011 I work as a full time researcher in the TSF group. I obtained a PhD degree in computer science from University of Rennes in November 2009. From January 2010 to September 2011, I worked as a Postdoc in the FG Inet group, Berlin. I defended my Habilitation a Diriger les Recherches in June 2019.

My current focus is on algorithmic (adversarial) transparency: how to infer properties of remote (online) algorithms? Which properties can be inferred at reasonable cost? Can such approaches be used by societies to dispute with tech giants over the control of our digital existences?

I'm broadly interested in algorithms and graphs. I try to apprehend graphs both as mathematical objects, and as models of the interaction structure of real world objects. I'm particularly interested in algorithms that rely on/exploit/capture such graphs, and how to tailor them for "typical" "real world" interaction structures, where the meaning of "typical" and "real world" is defined by a product of fashion, context, and mathematical docility. For this I typically try to find a blend of abstract models and data mining approaches.