University of Strathclyde Law School, Centre for Internet Law & Policy, Centre for the Study of Human Rights Law & Global Digital Human Rights Network (GDHRNet)

Symposium

Artificial Intelligence and Human Rights-Based Decision-Making

30 August 2023
Technology & Innovation Centre, 99 George Street, Glasgow G1 1RD
Executive Suite B

Zoom link: https://strath.zoom.us/j/87307245742 (password: 467306)

9.00 – 9.15 Welcome: Prof. Adelyn Wilson/Head of Strathclyde Law School; Prof. Mart Susi/GDHRNet Action Chair; Dr Birgit Schippers/Strathclyde Law School

9.15 – 10.45: Artificial Intelligence and Human Rights
Chair: Jukka Viljanen (Tampere University)

Sejla Maslo Cerkic (Organisation for Security and Co-operation in Europe, Mission to Bosnia and Herzegovina)
The impact of AI on the rights of journalists and human rights defenders - a human rights perspective

Vygantė Milašiūtė (Vilnius University)
The use of AI in courts and its impact on the right to fair trial

Philipp Mahlow (University of Innsbruck and University of Cologne)
'Where the hell did you get that idea, Hal?' Explainable AI-Systems: A Human Rights-Centric Approach

Mart Susi (Tallinn University)
Trust in AI from a human rights perspective

10.45 – 11.00: Comfort break

11.00 – 12.30: Towards Digital Human Rights?
Chair: Elaine Webster (University of Strathclyde)

Adam Harkens (University of Strathclyde)
Unravelling Assumed Discretion in Algorithmic Policing: A Defining Test for Human Rights Safeguarding in the UK

Francisco Bedecarratz Scholz (Universidad Autónoma de Chile)
Approaches to digital human rights in South America

Evelyne Tauchnitz (University of Lucerne)
Negotiating for Humanity: The Council of Europe's Framework Convention on Artificial Intelligence, Human Rights, Democracy and Rule of Law discussed from an ethical perspective

Jukka Viljanen (Tampere University)
The Council of Europe and the fight against disinformation

12.30 Lunch

ABSTRACTS AND PRESENTER BIOS

Panel 1: Artificial Intelligence and Human Rights

Sejla Maslo Cerkic (Organisation for Security and Co-operation in Europe, Mission to Bosnia and Herzegovina)
The impact of AI on the rights of journalists and human rights defenders - a human rights perspective

Bio:
Sejla Maslo Cerkic, PhD, Organisation for Security and Co-operation in Europe, Mission to Bosnia and Herzegovina. Working as a National Legal Officer in the Human Rights Section of the OSCE Mission to Bosnia and Herzegovina, previously an assistant professor at the Faculty of Law of the "Dzemal Bijedic" University of Mostar, Bosnia and Herzegovina, lecturer in contemporary challenges in international human rights law at Tallinn University School of Governance, Law and Society.

Abstract:
Artificial intelligence has a major impact on the work of journalists and human rights defenders. From obvious benefits which have made the journalistic work and the efforts of human rights defenders more efficient and impactful to serious concerns that human

journalists may be replaced by AI in the near future, the debate has become more relevant with the emergence and the wide use of new tools such as chat bots. Furthermore, the challenges posed by different surveillance technologies have made the work of journalists and human rights defenders more dangerous in the light of a global political crisis and unsafe working environments. Many of these concerns are related to the protection of freedom of expression, the right to privacy, issues with safety and labor rights, etc. The presentation will focus on the human rights implications of AI, as well as some liability concerns with regards to automatically created journalistic content. The impact of AI on investigative journalism and human rights will be presented, with an overview of the most relevant policies developed by the international human rights bodies on the subject.

Vygantė Milašiūtė (Vilnius University)
The use of AI in courts and its impact on the right to fair trial

Bio:
Vygantė Milašiūtė is an associate professor at the Faculty of Law, Vilnius University, and an Action Vice Chair of the GDHRNet. She was a civil servant at the Ministry of Justice of Lithuania and has experience of cooperation with the UN, Council of Europe, and EU institutions.

Abstract:
The presentation examines how the use of AI in courts can affect the exercise of the right to fair trial. It provides some examples of the use of AI for judicial tasks. It then looks into issues of access to justice, institutional requirements for a tribunal (impartiality, independence, establishment by law), the right to adversarial proceedings and equality of arms, the presumption of innocence, the protection against entrapment, the rights of defence, the right to a reasoned decision and the requirement of the overall fairness of trial. Inter alia, it addresses the issue of whether there is a human right to a natural judge. It looks at the possibilities of improving the law and lawyers, improving AI, and the possible necessity to acknowledge certain limits of how far the improvement can go. It shows why collaboration between a judge and a machine is the preferred solution.

Philipp Mahlow (University of Innsbruck and University of Cologne)
'Where the hell did you get that idea, Hal?' Explainable AI-Systems: A Human Rights-Centric Approach

Bio:
Philipp Mahlow works at the University of Cologne in a research project exploring training data quality requirements for AI systems ("KITQAR") and is a PhD student at the University of Innsbruck. His thesis explores legal explainability requirements for AI systems in current regulatory acts of the European Union.

Abstract:
The insufficient interpretability of AI systems and their impact on human rights is one of the main inhibitors for the large-scale use of AI systems. If the affected person cannot comprehend a decision, it is more difficult for him or her to accept it and trust the system producing that decision. Many recent examples show that decisions of modern and complex algorithms cannot be understood even by their deployers and that this so-called black-box problem can cause great damage. Computer science reacted to this with the development of so-called Explainable AI approaches (in short: "XAI"). The necessity of explainability requirements is also justifiable from a legal point of view due to the threats to human rights and fundamental principles of the rule of law. Thus, it is not surprising that several European Digital regulatory acts now consider explainability requirements in different capacities. However, these are not without problems. Their wording misleads about the actual state of the art in the XAI field. At the current time, counterfactual approaches in particular should be used to enable the explanation of a system decision in individual cases.

Mart Susi (Tallinn University)
Trust in AI from a human rights perspective

Bio:
Doctor Juris Mart Susi is the Action Chair of the Global Digital Human Rights Network. Mart Susi has an MA degree in Sociology from the University of Wisconsin-Madison (USA) and a Doctor Juris degree from the University of Tartu (Estonia). He has held senior positions in several academic institutions. Currently, he is heading the law program at Tallinn University and has the position of Professor of Human Rights Law. Mart Susi has initiated and is leading several research and development projects funded by the European Commission and the Nordic Council of Ministers. He has edited and is currently editing several research books on the topics of new human rights, the digital dimension of human rights, and the controversy around the meaning of human rights. He is developing the Internet Balancing Formula and has lectured on the topic at various universities in Europe, Asia, and South America. He is also involved as an expert for the European Commission and non-governmental organizations.

Abstract:
The proposition of turning unfair normativity into fair judicial or quasi-judicial outcome cannot be validated for the digital domain, where automated solutions or the artificial intelligence are allowed to make decisions. This is because of two reasons. The first is because of a phenomenon which is primarily intuitionistic and asserts that the concept of trust in the digital domain has a blind eye towards the question of normative fairness. This is related to the objection from Radbruch's disavowal formula. Digital domain operators – automated systems and the artificial intelligence, take for granted the positive features of positive or quasi-positive rules. Judges refusing to apply a legal norm which would lead to extremely unfair outcome do so because they can exercise

distrust towards law. Artificial intelligence and automated systems, on the other hand, are designed to trust law. To put it figuratively, the degree of our trust in physical judges is related to our expectation that they are capable of distrusting law. And our trust towards automated systems and artificial intelligence is weaker because we assume that they completely trust law. This is the context which will be discussed in the lecture.

Panel 2: Towards Digital Human Rights?

Adam Harkens (University of Strathclyde)
Unravelling Assumed Discretion in Algorithmic Policing: A Defining Test for Human Rights Safeguarding in the UK

Bio:
Adam is a Lecturer in Public Law at the University of Strathclyde, and member of the Strathclyde Centre for Doctoral Training in Human Rights-based Decision Making. Previously, Adam was a Research Fellow at the University of Birmingham, and a Research Assistant at King's College London and the Public Law Project. He holds a PhD in Law from Queen's University Belfast.

Abstract:
Algorithmic policing – i.e., the use of software algorithms to process digital input data, to generate outputs capable of supporting or triggering police decisions - is proliferating in the UK. Amidst political and economic pressures, police are actively seeking digital solutions to assist in the pursuance of their public duties and obligations to protect the public and/or to prevent and investigate criminal offending. Concerns have been raised about potential human rights interferences caused by algorithmic policing practices, exacerbated by technical opacity and the potential for discrimination and automation bias. There are however even more fundamental and unanswered questions which derive primarily from the allocation and exercise of discretionary policing powers to (or by) human actors, as opposed to problems inherent to digital technology: from where does the legal authority to deploy algorithmic policing tools derive? With whom can the lawful power to determine when, and how, to use these tools be said to reside? And what ought the appropriate legal limits of this power be? Using FOI data, this paper demonstrates a clear accountability gap surrounding algorithmic policing in the UK. In 'rights-critical' contexts, algorithmic policing has the potential to displace human discretion and/or introduce new forms of discretion to police decision-making in unpredictable ways. Recognising that discretionary power in this context is poorly understood, it argues for the need to unravel both, firstly, how and why these tools are being used; and secondly, how the police have assumed discretionary legal authority to make use of these tools, and how they define the boundaries of this power. Only then can we begin to identify appropriate legal limits that will robustly safeguard human rights for individuals affected by algorithmic policing decisions.

Francisco Bedecarratz Scholz (Universidad Autónoma de Chile)
Approaches to digital human rights in South America

Bio:
Dr. Francisco Bedecarratz Scholz is an assistant professor of Criminal Law and director of the "Artificial Intelligence and Law" Research Group at Universidad Autónoma de Chile. He holds a PhD in Law from Philipps-Universität Marburg, Germany, and an LLM from the same institution. His research centers on corporate criminal liability, artificial intelligence, and cybersecurity.

Abstract:
This presentation explores the diverse approaches to digital human rights in South America. It delves into the region's unique challenges, advancements and initiatives aimed at recognizing and safeguarding human rights in the digital sphere. By examining legal frameworks, policy developments and judicial precedents, this presentation highlights the efforts made by South American countries to protect individual's rights to privacy, freedom of expression and access to information in the digital age, among others. Furthermore, the presentation sheds light on the progress achieved, identifies existing gaps and offers insights into potential avenues for further enhancements in the field.

Evelyne Tauchnitz (University of Lucerne)
Negotiating for Humanity: The Council of Europe's Framework Convention on Artificial Intelligence, Human Rights, Democracy and Rule of Law discussed from an ethical perspective

Bio:
Dr. Evelyne Tauchnitz conducts research on the digital transformation, ethics, human rights and peace at the Institute of Social Ethics ISE, University of Lucerne, Switzerland. Evelyne is also a Senior Fellow at the Centre for International Governance Innovation and the Centre for Technology and Global Affairs at Oxford University.

Abstract:
The Council of Europe's Draft Convention on Artificial Intelligence (AI) aims to establish a comprehensive framework for regulating AI. The Convention's objective is to ensure that AI is developed and used in a manner that promotes and protects human rights, democracy, and the rule of law in the digital environment. It aims to set forth principles that guide the responsible design, development, and deployment of AI systems to avoid potential harms and negative impacts on individuals and societies. The Convention is intended to be a global instrument that is no only limited to member states, but attracts participation from states around the world who align with the values of human rights, democracy, and the rule of law. Given its global ambition, it is key to ask on what ethical principles and values the Draft Convention builds and in how far it provides the necessary tools and mechanisms to translate these ethical principles into practice. Will

it make a difference in the real world? Are its legal provisions at the same time open enough to encourage wide participation of states around the world, while at the same time being also able to safeguard individuals and groups' rights and freedoms against abusive AI systems? Is there a trade-off between the political encouragement of wide adoption of the Draft Convention and the strength of its protective legal provisions?

Jukka Viljanen (Tampere University)
The Council of Europe and the fight against disinformation

Bio:
Dr Jukka Viljanen is Professor of Public Law, Tampere University, Finland. He is a Member of Finnish Human Rights Delegation (2020-2024) and an expert before the Constitutional Law Committee of Finnish Parliament. His recent research projects are ALL-YOUTH (2018-2023) focusing on youth participation in law drafting and EVOLUTIVE (2021-22) focusing on monitoring and implementing Council of Europe Human Rights Treaties. Professor Viljanen has published extensively on the European Court of Human Rights. He was one of the authors and editors in Human Rights Law and Regulating Freedom of Expression in New Media, Lessons from Nordic Approaches published with Routledge.

Abstract:
The disinformation is becoming more sophisticated, meaning that it is more difficult to recognize from real news sources and judge the credibility of the message. This is not only done by individuals or small groups, but even states and state-owned media companies are behind disinformation campaigns. The ECtHR's approach to disinformation has been under the scope of freedom of expression and more specifically in relation to media pluralism and hate speech. The pluralistic media and ensuring quality media environment are key elements to combat disinformation. The hate speech brings into discussion the liability of not just those expressing the disinformation and conspiracy theories, but also those distributing it via social media platforms. It is also relevant to understand context and impact of disinformation. The context of elections was discussed in the Grand Chamber case of Sanchez v. France, but it also raised questions on the limits of a general political debate. Finding a fair balance is critical countering disinformation. The "disinformation or propaganda" argument might be used in order to attack political opposition or particular minority groups e.g. LGBTQ-community. The connection of disinformation to other rights than freedom of expression should be also acknowledged. Often the current disinformation is falling into the scope of Article 8, right to private life, including protection of reputation. One of the most notable connections of disinformation and conspiracy theories is related to anti-vaccination campaigns which was emphasized during the Covid 19 pandemic. Disinformation often targets also certain vulnerable groups whether ethnic minorities or refugees, so there is a connection also to Art 14 and Protocol 12, ensuring non-discrimination.