

SAAREMAA TRAINING SCHOOL in human rights and the AI GDHRNet 06 – 08 July 2023



Global Digital Human Rights Network (GDHRNet), EU COST Action no CA19143 (EU Cooperation in Science and Technology Program) organizes a training school “Human Rights and the Artificial Intelligence”. The training school will be organized at Georg Ots Spa Hotel in Kuressaare, the island of Saaremaa, Estonia: <https://www.gospa.ee/eng>

The training school will open to the participants theoretical aspects and practical challenges related to the fast development of the AI and automated decision-making systems from human rights perspective. Beyond-state-of-the-art views will be discussed by scholars whose research mainly focuses on digital human rights. Participants will receive a certificate of attendance from the GDHRNet.

The training school covers the **following costs** for participants attending in person: travel expenses to Kuressaare, plus daily allowance of 192 euro per day for four days (the day of travel on 05 July and three days of the training school), so in total 768 euro. Lectures will be each day from 10.00 to 17.00 with breaks. Reimbursement of the costs will be paid within one week after submission of travel expense documents, signatures at attendance sheets for each day will be required. Accommodation has to be covered from the daily allowance.

The training will be in hybrid format. Preference is given to students attending in person, the limit of students will be 25.

The training school is open for Master and PhD students and professionals who are interested in digital and/or human rights. Bachelor level students may be considered on the basis of expression of interest. **To apply, please write by 15 May 2023 to professor Mart Susi at: mart.susi@tlu.ee** Please submit a short (maximum 1 page) statement of your interest in the subject and short cv (including your educational background). **Please also indicate whether you plan to attend in person or online.** Applications are reviewed and approved as they arrive, so the earlier you apply the earlier you will receive the decision.

The program and speakers

Wolfgang Benedek (Institute of International Law and European Training and Research Centre for Human Rights and Democracy of University of Graz)

The debate on new digital rights

Several recent initiatives have proposed new human rights in the digital sphere. They are to respond to new challenges to human rights in an increasingly digitalized world. However, is there really a need for genuine new digital human rights or would it suffice to develop existing rights by interpretation to deal with the new threats from cyberspace? As the proposals claim

the emergence of new principles and rights, in which stage is the development of new digital human rights? Can all digital rights proposed also be considered as human rights? How have European institutions reacted to the proposals or which regulatory efforts have they undertaken? If the new rights are proposed to be protected at the European level, what about the universal level of a global cyberspace? Specific attention will be given to the regulation of artificial intelligence from a human rights perspective.

Lana Bubalo (Associate professor of law at the University of Stavanger School of Business in Norway)

Legal protection against discrimination by AI

The use of AI is inescapable in our modern society. It has many benefits, but it also raises many concerns. Even though AI can be more objective than humans, in reality algorithms make decisions based on human input, which can be bias and exclusionary. AI used for automated decision making can lead to discrimination based on race, age and gender undermining equal opportunities and causing serious social consequences, particularly for vulnerable groups. This is why discrimination by AI has been identified as serious issue and challenge that is in the forefront of legal discussions on AI. In this lecture, we will explain what makes AI discrimination different from other types of discrimination and whether the current legal framework is enough to ensure efficient protection of the right to non- discrimination. Real life examples of AI discrimination (in health care, hiring process and criminal justice) will be used as a basis for in- class discussions in order to try to uncover answers to many questions the use of new technology poses for fundamental human rights, such as lack of algorithmic transparency, contestability, accountability and personhood issues.

Sejla Maslo Cerkić (Organisation for Security and Co-operation in Europe, "Dzermal Bijedic" University of Mostar, Faculty of Law)

Artificial intelligence and its impact on journalists and human rights defenders – the human rights perspective

Artificial intelligence has a major impact on the work of journalist 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 dangerous 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. After a general introduction on the risks and benefits of AI for the journalism and the media in a contemporary setting, the lecture 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.

Matthias Kettemann (Leibniz Institute for Media Research, Research Program Head; University of Innsbruck, Department of Legal Theory and Future of Law, Professor of Innovation, Theory and Philosophy of Law - Head of Department)

Beyond the "Black Box": potentials and limits of explaining AI-based recommendations and decisions in the Digital Services and the AI Act

The presentation explores the growing need for transparency and explainability in artificial intelligence (AI) systems used in digital services, especially recommender systems. The lecture will discuss the benefits and limitations of explainability, along with the regulatory framework provided by the DSA and AI Acts. Highlighting how AI systems often operate as a "black box," making decisions that are difficult to understand or explain, the lecture will show how the lack of transparency can lead to issues such as bias and discrimination. The lecture will also show how the DSA and AI Act approach "explainable AI" and which regulatory issues persist - and how to overcome them, including through ethical approaches to AI that go beyond regulatory obligations.

Artūrs Kučs (Judge at the Constitutional Court of Latvia, associate professor of Latvia University)

3rd part liability. Does the Delfi principle still stand?

The lecture will explore the broader context how the Internet has changed the ways we receive and impart information, the notion of media and the way media operate. What are developments as regards intermediary liability from the EU and the CoE? Some key ECtHR judgments will be explored, in particular after the Delfi vs. Estonia judgment: Index vs. Hungary, Pihl vs. Sweden, Hoinez vs. Norway. The lecture will also explore developments and discussions about intermediary liability in the US?

Mart Susi (Professor of Human Rights Law at Tallinn University, Chair of Global Digital Human Rights Network)

AI and trust from the human rights perspective

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.

Jukka Viljanen (Professor of Public Law at the Tampere University, Finland)

AI and freedom of expression: European perspective

Discussion on AI and Freedom of expression has arisen from several perspectives. There have been raised questions related to role of search engines, algorithms polarizing political discourse, even first defamation cases against Chat GPT are now before courts. There is also discussion of positive sides of using AI e.g. combating hate speech. In this lecture, we focus on some of the main issues. How should we review the freedom of expression doctrine in

light of AI? What are the key elements of the prevailing freedom of expression doctrines that needs to be developed due to AI? What are the roles and responsibilities of different actors (authorities, tech companies, individuals) in the age of AI?