

The purpose of the project is to create a universal model for non-discrimination in algorithmic decision-making. In order to answer the main research question: “which grounds of algorithmic decision-making may violate the universal prohibition of discrimination?”, the project’s hypothesis assumes that universal grounds of non-discrimination may be violated when Artificial Intelligence’s decision-making process is based on such grounds like race, sex, language or religion. International human rights law will be studied in order to determine the role of non-discrimination in algorithmic decision-making.

Context: The technological progress has increased the interest of States in the deployment of AI and robotics, while fulfilling their functions, consequently affecting human rights. Such a phenomenon is related to the growing role of algorithmic decision-making in civil service, particularly in courts’ rulings. The U.S. judiciary system, for instance, uses predictive algorithms to try accused in criminal proceedings. This raised the question on whether algorithms should mathematically arbitrate fairness for human beings. The problem is that the system has been created to make “risk-scores” on the prediction about groups of people, who share similar features, but not about the individuals themselves. When the algorithmic biases will be taken into account, the problem with the prohibition of discrimination arises. Another example is a software called *Predictive Policing* designed to predict parts of the city that are more prone to criminal appearance and to locate police resources respectively. One can ask on its relationship with the prohibition of discrimination based on race, nationality, ethnicity, gender, age or any other ground. It might be a case when the police resources are over-located due to the fact that a specific group (national, racial, ethnical or other) lives in the area. *ROBORDER* platform, on the other hand, aims at developing and demonstrating a fully-functional autonomous border surveillance system with unmanned mobile robots of all types (aerial, water surface, underwater and ground vehicles), in order to detect border crossing of irregular migrants and of environmental pollution. Although intelligent robot can serve for acceptable surveillance, for the time being nobody can predict its weapon-attached use. This leads to the last but not least example of AI robots, namely autonomous weapon systems. They can be used both during the time of peace and armed conflict, thus transferring the license for killing from a human being to an intelligent robot.

All of these examples raise not only ethical doubts, but also legal questions. Many objections are related to the possible lack of neutrality of algorithms (often collided with black boxes), the acceptable level of human control over AI’s performance and the responsibility for the violations of the prohibition of discrimination. This is the reason why it is necessary to develop the cooperation between lawyers as well as producers and their contractors. This is also why the project can become useful for creating an effective balance between State and individual interests. The project deals with substantive and procedural guarantees of compliance with human rights, particularly the potential of Universal Periodic Review for naming and shaming violations of the universal grounds of the prohibition of discrimination, resulting from the use of algorithmic decision-making.

Consequently, the project will seek for the elements of responsibility for human rights violations. In this respect, it is necessary to examine the emerging components of international robotic law, namely the principle of meaningful human control on the AI-based system, the principle of robot obedience with human orders, the obligation to keep constant and special control while using critical systems (such as autonomous vehicles or autonomous weapon systems). Activities undertaken in this purpose will be multi-dimensional and examine obligations of both States and AI industry, developing AI and robotics. Taking into account the impasse in the discourse on potential subjectivity of robots and their responsibility, the consequence shall be the development of security and responsibility standards (or regulations) with regard to the violations of international human rights law. Technologies created by humans shall serve for the humanity, whilst not posing a threat to it.