What Ways to Program Autonomous Systems (such as Self-Driving Cars) are Permissible in a Liberal-Democratic *Rechtsstaat*?

Special workshop at the IVR World Congress 2019 in Lucerne
Friday, July 12th 2019, 13:30 to 19:00, University of Lucerne

13:30 - 14:30	Prof. Dr. Christoph Lütge (TU Munich) *Ethics and Innovation in Artificial Intelligence*
14:30 - 15:00	Andrej Dameski (University of Luxembourg) *The Hero in the Machine: Virtue Ethics in the World of AI*
15:00 - 15:30	Antonio Bikić (LMU Munich) *On the Ontological and Moral Status of Artificial Agents*
16:00 - 17:00	Prof. Dr. Annette Dufner (University of Bonn) *On Saving the Greater Number in Rescue Conflicts*
17:00 - 17:45:	Anja Bodenschatz (TU Munich/University of Cologne), Dr. Matthias Uhl (TU Munich), Dr. Gari Walkowitz (TU Munich) *Managing Ethical Dilemmas in Human-Machine Interactions through Randomization*
18:00 - 18:30	Adriano Mannino (University of Bern) *May Machines Sometimes Do What Humans Are Not Permitted To Do? Remarks on Interpersonal Aggregation and Fair Lotteries in AI Contexts*
18:30 - 19:00	Marina Moreno (University of Bern) *Aggregate Relevant Interests Only? Theories of "Limited Aggregation" and the Problems They Face*

Description / Call for Papers:

What Ways to Program Autonomous Systems (such as Self-Driving Cars) are Permissible in a Liberal-Democratic *Rechtsstaat*?

Suppose you're steering a car and facing the dilemmatic choice of either killing two pedestrians or one biker – it's no longer possible to avoid human casualties. Should you aggregate the victim counts and save the greater number, i.e. kill the biker? Or should you (time allowing) randomise your decision, i.e. toss a fair coin or perform a fair lottery, so as to allocate every person involved an equal 50% survival chance? Or would fairness require a weighted lottery, i.e. allocating a 1/3 survival probability to the biker and a 2/3 probability to the pedestrians?

Slow human reaction times usually save us from having to make such life-and-death decisions when accidents occur on the road. However, the algorithms steering self-driving cars are likely to become able to process the relevant data sufficiently quickly. Since these algorithms are our creation, we face a large class of *ex ante* life-and-death decisions when programming them. This raises a number of moral and legal issues. Among them is the question of the permissibility of aggregating lives: If we program a self-driving car to automatically save the two pedestrians at the expense of the one biker in the above case, are we violating the biker's dignity or autonomy? And would it therefore be incompatible with the principles of the liberal *Rechtsstaat* to allow democratic majorities to opt for an aggregative algorithm?

Some philosophers and legal scholars have argued that because we are in a "veiled" *ex ante* situation when programming autonomous cars, i.e. because we are ignorant about the exact place we will occupy in potential future accidents, aggregation may be justifiable to each person and permissible on liberal grounds after all: *Ex ante*, each person's chance to end up among the two pedestrians is twice as high as their chance to be the biker; hence, each person benefits in expectation from the two pedestrians being saved. – But is this reasoning sound? And might it have worrisome implications with regard to the permissibility of, say, using military force against a civilian aircraft hijacked by terrorists, which Germany's *Bundesverfassungsgericht* has declared unconstitutional based on Articles 1 and 2 of the *Grundgesetz*?

The above-mentioned questions will constitute a main focus of the workshop. However, we highly welcome any contributions addressing moral-cum-legal issues concerning AI and autonomous systems (not limited to cars) in the context of the liberal-democratic *Rechtsstaat*.