

2021-006 vom 26.01.2021

VolkswagenStiftung fördert gemeinsames Projekt mit Saarland-Uni JProf. Schmidt von der TU Dortmund forscht zu Entscheidungen von Künstlicher Intelligenz

Wie können wir sicherstellen, dass sich Maschinen moralisch verhalten? Das ist eine Fragestellung, der sich JProf. Eva Schmidt von der TU Dortmund in einem gemeinsamen Projekt mit der Universität des Saarlandes widmet. Die VolkswagenStiftung fördert die Forschung von Schmidt und einer Doktorandin mit rund 160.000 Euro.

Big Data, Digitalisierung, Deep Learning – neue Technologien werfen neue Fragen auf. Wie soll sich beispielsweise ein autonom fahrendes Auto „entscheiden“, wenn es in einer Zwangslage vor der Wahl steht, einen Menschen zu überfahren? Soll es eher den Rentner als den Jugendlichen treffen? Der eine hat sein Leben gelebt, der andere hat es noch vor sich. Welche Entscheidung ist in einer solchen Situation die richtige für die Maschine?

Dies ist eine Frage im Spannungsfeld Künstlicher Intelligenz (KI) und Gesellschaft. Die VolkswagenStiftung hat insgesamt rund 12 Mio. Euro in ihrer Förderinitiative „Künstliche Intelligenz – ihre Auswirkungen auf die Gesellschaft von morgen“ bewilligt, damit fach- und länderübergreifende Forschung zur verantwortungsvollen Weiterentwicklung von KI-Systemen betrieben wird. Damit unterstützt die Stiftung acht interdisziplinäre und internationale Forschungsverbünde aus den Gesellschafts- und Technikwissenschaften. Wissenschaftlerinnen und Wissenschaftler etwa aus den Rechts-, Sprach- und Sozialwissenschaften sowie aus Informatik, Medizin, Philosophie und Cybersicherheit haben sich in dem Projekt zusammengefunden. Die Vorhaben sind auf drei bis vier Jahre angelegt, Start in Dortmund soll im späten Frühjahr sein.

JProf. Eva Schmidt vom Institut für Philosophie und Politikwissenschaft der Fakultät Humanwissenschaften und Theologie ist dabei. Sie forscht schon länger zu diesem Themenkreis: Welche Chancen bietet KI? Wo liegen Risiken? Und konkreter: Wieso ist es wichtig, dass wir die Entscheidungen von KI-Systemen verstehen? Wie fördert die Erklärbarkeit von KI-Systemen z. B. Vertrauen und Verantwortung?

In Schmidts Projekt „Explainable Intelligent Systems (EIS)“ arbeiten die Disziplinen Informatik, Philosophie, Psychologie und Recht zusammen. Im Projekt geht es um die Erklärbarkeit von KI-basierten Entscheidungen und damit um eine der Schlüsselfragen des Einsatzes von KI-Systemen in der Gesellschaft. Projektbeteiligte sind neben JProf. Schmidt von der TU Dortmund JProf. Lena Kästner, Prof. Georg Borges, Prof. Ulla Wessels, Dr. Markus Langer und Prof. Holger Hermanns, alle Universität des Saarlandes in Saarbrücken. „Unter anderem befassen wir uns mit der übergreifenden

Fragestellung: Wie können und sollten intelligente Systeme so konzipiert werden, dass sie erklärbare Empfehlungen geben?“, so JProf. Schmidt. An der TU Dortmund wird sie von Sara Mann unterstützt, die im Rahmen des Projekts promovieren wird.

Im Rahmen ihrer Forschung zu erklärbarer KI hat Schmidt auch in Hannover, Zürich und Cambridge vorgetragen und zuletzt online auf der *Usability in Germany*-Tagung zum Thema „Autonomie und Gerechtigkeit im Rahmen Künstlicher Intelligenz“ referiert. In Kooperation mit dem *Leverhulme Centre for the Future of Intelligence* in Cambridge, der TU Delft und der Leibniz Universität Hannover richtet das EIS-Projekt die Workshop-Reihe „Issues in Explainable AI“ aus. Der Dortmunder Workshop der Reihe ist für den Herbst 2022 geplant.

Weiterführende Informationen:

<https://explainable-intelligent.systems>

Bildinformation: Eva Schmidt ist Juniorprofessorin für Theoretische Philosophie an der TU Dortmund. Foto: Privat

Ansprechpartnerin für Rückfragen:

JProf. Eva Muriel Schmidt

Fakultät Humanwissenschaften und Theologie

Tel.: 0231-755 2835

E-Mail: eva.schmidt@tu-dortmund.de

Die Technische Universität Dortmund hat seit ihrer Gründung vor 52 Jahren ein besonderes Profil gewonnen, mit 17 Fakultäten in Natur- und Ingenieurwissenschaften, Gesellschafts- und Kulturwissenschaften. Die Universität zählt rund 33.440 Studierende und 6.500 Mitarbeiterinnen und Mitarbeiter, darunter etwa 300 Professorinnen und Professoren. Das Lehrangebot umfasst rund 80 Studiengänge. In der Forschung ist die TU Dortmund in vier Profildbereichen besonders stark aufgestellt: (1) Material, Produktionstechnologie und Logistik, (2) Chemische Biologie, Wirkstoffe und Verfahrenstechnik, (3) Datenanalyse, Modellbildung und Simulation sowie (4) Bildung, Schule und Inklusion. Bis zu ihrem 50. Geburtstag belegte die TU Dortmund beim QS-Ranking „Top 50 under 50“ Rang drei der bundesdeutschen Neugründungen.