Lecture series on the subject "The interplay between fundamental computer science and data science"
Chair: Prof. Dr. Florin Manea

Abstract and brief biography

4. Dirk Nowotka: On the robustness of neural networks

Abstract: Adversarial examples are a well-known phenomenon in the field of machine learning. They are an outstanding example of the (perceived) fragility of those methods. We are going to consider techniques of how to interpret and defend against those examples. We will conclude with questions for fundamental studies in computer science.

Brief Bio:
Dr. Dirk Nowotka leads the Dependable Systems group of the Computer Science department at Kiel University, Germany. Prior to joining Kiel as a Heisenberg-Professor in 2011, he was a research scientist at the Stuttgart University (2004-2011), Germany, where he gained his habilitation, and the ETH Zürich (2004), Switzerland. He completed his PhD in mathematics from the University of Turku (2004), Finland. Dirk's primary field of research is the theory and practice of automated mathematical and logical procedures for the safety and security analysis of software systems. One particular research interest of him is safety in the field of artificial intelligence.