

Preface

Just today, as we are making the final corrections to the present volume, an online article rolls into our newsfeeds. There we read that in June of this year, Google programmer Blake Lemoine went public. He was concerned with the 'Language Model for Dialogue Applications' (LaMDA), an intelligent speech bot from Google specialized in dialogues. "As first reported in the Washington Post, programmer Lemoine, after his conversations with LaMDA, no longer believes that this machine is merely a tool. He demands that LaMDA be treated as a person. He believes LaMDA has gained consciousness." – "Can this be?" asks the author of the article. – "Short answer: most likely not."¹

And yet, no other topic in recent years has triggered such a storm of enthusiasm and simultaneously such a wave of uncertainty: artificial intelligence. Intelligent software will automate work, understand images and text and make medical diagnoses. It enables cars to drive independently and supports scientific work. On the basis of current innovations, a global revolution in business and industry can be predicted, which is being driven along ever further by the networking of 'smart machines'. At the same time, artificial intelligence promises to be able to analyze the masses of data produced in the course of digitalisation for patterns and to make automated decisions based on these, which will not only rival the quality of human analysis, but will be far superior to human work in terms of efficiency.

The current central paradigm of AI is machine learning, whereby deep learning seems to represent the most promising technological approach: software is no longer programmed for its specific tasks but is trained on

1 "Künstliche Intelligenz, [...] was kommt da auf uns zu?" October 11, 2002. URL: https://krautreporter.de/4601-kunstliche-intelligenz-so-verstandlich-wie-moglich-erklart?utm_campaign=pocket-visitor, translated from German.

large amounts of data – on Big Data. Here, however, the limits of the new technology begin to become apparent, as data volumes are not always available in the desired quality and quantity. It also requires basic or retrospective labelling by humans as well as the processing of data by means of symbolic AI procedures, therefore by supporting Small Data approaches. It is not uncommon for social discrimination to be inherent in the data sets, to which human preselection, categorisation and corresponding annotation make a decisive contribution. Inevitably, this leads to the integration of cognitive biases into a given AI: prejudices embedded in society are reproduced and to some extent further reinforced. Moreover, if machine decisions are made on the basis of deep learning systems, the decisions made by the machine are also non-transparent for humans. These are quite rightly characterised as black box procedures, which are accompanied by considerable legal problems. And finally, but importantly, the fundamental question arises as to whether people can trust systems with which they are supposed to collaborate, when not even the systems' very creators can reconstruct and interpret the decisions of their technical creatures.

The volume is therefore intended to explicitly show the limits of AI, to describe the necessary conditions for the functionality of AI, to reveal its attendant technical and social problems, and to present some existing and potential solutions. At the same time, however, we also want to describe the societal and attending economic hopes and fears, utopias and dystopias that are associated with the current and future development of AI. This has led us to gather in this volume authors and contributions from a variety of scientific disciplines, meaning that computer scientists, engineers, mathematicians, as well as media studies experts, social scientists and economists have their say on an equal footing. And lastly, but foremost, we thank our English-language editor, Richard Slipp, who has repeatedly defied DeepL and all other artificially intelligent translators and pointed out their limitations, thus making the volume in its present form possible.

Peter Klimczak and Christer Petersen
October 2022