How supercomputers make artificial intelligence

Prof. Dr. Dr. Thomas Lippert, Forschungszentrum Jülich, Institute for Advanced Simulation and Jülich Supercomputing Centre

The triumphant advance of large artificial neural networks, especially large basic models such as Large Language Models (LLM), was only made possible by the rapid development of supercomputers of the Tera, Peta and now Exascale class, which has been taking place since the beginning of the millennium. You may be asking yourself whether it is not possible to achieve the goal with less massive and faster computing systems, perhaps just a little slower? In my presentation, I would like to make clear the enormous effort in terms of computing power and data throughput, but also energy and time, that is actually required to train today's largest basic models. It is not only the huge quantity of data that plays a decisive role here, but also the necessary quality of the data. Such basic models have the potential to fundamentally change basic and applied research in almost all areas of research and development, far beyond a national perspective.