[The original, German version of this article appeared on 14 December 2020 in Frankfurter Allgemeine Zeitung, Nr. 291, S. 20.]

[Translated with www.DeepL.com/Translator (free version) - AI technology made in Europe, see https://en.wikipedia.org/wiki/DeepL_Translator - and subsequently polished by the authors.]

The third wave of artificial intelligence

Why just now this key technology offers a historic opportunity for Germany and Europe

By Holger Hoos and Kristian Kersting

On the one hand, humanity is facing extraordinary challenges - arising from climate change, pandemics, geopolitical shifts and demographics. On the other hand, there has also been enormous progress: the molecular scalpel, CRISPR-Cas9, is revolutionising precision medicine, gravitational waves have been detected, and cheaper, reusable space vehicles are providing previously unimaginable access to space. This is both gratifying and necessary, because the major problems of our time demand solutions that go far beyond what is currently feasible, both scientifically and technically. Artificial intelligence (AI) has a special role to play in this context, for two reasons: Firstly, these problems are caused and exacerbated, at least in part, by natural limitations of human intelligence. Secondly, as the next stage of the digital transformation, AI is a broadly applicable general-purpose technology with a wide range of applications.

It is worrisome that there is still confusion about what AI actually is. At times, it is used to refer to the vision of machines that exhibit the full spectrum of human intelligence and can thus replace or surpass humans, at least in principle, with understandably worrying consequences. On the other hand, for some time now, there has been a tendency to equate AI with machine learning or, even more narrowly, with so-called deep learning using artificial neural networks. Both of these notions of AI are misleading.

The term "artificial intelligence" (AI) was coined by American computer scientist John McCarthy in 1956, and has since been defined as the quest for computer programmes that reproduce intelligent (not necessarily human) behaviour. The central question of AI, however, was posed as early as 1950 by the British pioneer of computing, Alan Turing: Can machines think?

Indeed, there has recently been impressive and important progress in AI, especially in the field of machine learning, in terms of basic research as well as in applications. So-called transformers can add to or almost completely compose texts based on minimal input. The programme "AlphaFold 2.0", recently developed by Deep Mind, has proven capable of predicting the three-dimensional structure of proteins with the accuracy of laboratory experiments, and can thus help to better diagnose and cure diseases, or to design enzymes specifically for energy production or to break down pollutants. We are (fortunately) very far from being able to realise general AI, covering the full range of human intelligence. At the same time, current AI technology encompasses much more than machine learning. In addition to learning, approaches such as logical and mathematical reasoning, modelling of knowledge and knowledge processing, optimisation of processes and decisions, and dealing with multiple, often conflicting objectives also play an important role - which, of course, is also the case in the context of human intelligence.

As an example, consider the following situation: You are driving over land in a car and see a traffic sign with a speed limit of 120 km/h at the entrance to a town. What would we do? Step on the accelerator and push the speed limit to the limit? Probably not - because we know that only 50 km/h is allowed in built-up areas. But what would a modern AI system do? Using deep learning, it would be able to recognise the traffic sign and interpret it correctly - it would be obvious to the system that a speed limit of 120 km/h applies here. It has been trained for this beforehand. It would never realise by itself that this could not be right in this case. Humans, on the other hand, would recognise the 120 km/h speed limit as dangerous and drive at 50 km/h.

To get beyond narrowly limited capabilities, AI systems lack an crucial human ability: the sense of context that we take so much for granted in everyday life. We don't just solve a specific task. We acquire general knowledge and everyday skills, which we combine flexibly to master tasks and gain new insights. In this way, we grasp the meaning of a new word from just a few examples and avoid incorrect conclusions. The whale is not a fish, but a mammal. The perceived speed limit of 120 km/h is dangerous - we prefer to drive 50 km/h. And if we really get stuck, we can ask others for help and describe our problem to them. Humans are communicative, interactive and think in larger contexts.

It is precisely this level-headedness and flexibility that are essential elements of the "third wave of AI", which aims to produce AI systems with human-like communication skills and cognitive abilities that recognise new situations and adapt to them autonomously. During the first wave of AI, from 1956 to the 1980s, intelligent machine behaviour was still programmed by humans; computers could draw conclusions using a large set of if-then relationships using a well-specified program logic. A well-known example of such rule-based systems was the chatbot ELIZA by computer scientist Joseph Weizenbaum in 1966: this was the first computer program that could conduct conversations in English. The second wave of AI started in the 1980s and continues until today: Computers learn from examples, from data. Humans no longer have to laboriously cover all eventualities by means of countless, manually specified rules. Only the learning algorithm is programmed, the rest amounts to collecting examples of intelligent behaviour.

It is interesting to note that machine learning is not really new. Its origins can be traced back to the 1940s, 1950s and 1960s - for example, the simple model of a neuron devised by Warren McCulloch and Walter Pitts, which does not accurately describe the brain but became an important impetus for deep learning, or Arthur Samuel's adaptive program for playing checkers. However, only the advent of special processors, so-called GPUs, and the compilation of large amounts of data, such as Fei-Fei Li's ImageNet, made it possible, from around 2010 onwards, to train large artificial neural networks with millions of examples. This in turn led to well-known breakthroughs in image recognition or in playing board games, such as chess and Go. Today, everyone benefits from these developments - for example, through speech recognition in smartphones, or in cars, where driving assistants are increasingly providing support.

But these systems lack the ability to recognise and leverage larger contexts, especially when they have to work with little or unreliable data. And this is precisely where current developments come into play. The third wave of AI builds on the successes and concepts of the first two waves. Low-level perception (traffic sign recognition) and higher-level abstraction thinking (traffic rules) are being merged, so that conclusions and decisions can be contextualised (driving in a town and no longer on the highway) and communicated in a human-like way (asking the passenger). For humans, Nobel laureate Daniel Kahneman contrasts, in his book "Thinking, Fast and Slow", the fast, instinctive system (neural networks and probability models) and the slower, methodical, analytical system (logic).

A simple approach to this is to embed learning procedures into systems that guarantee, for example, through provably correct logical rules of reasoning and through suitably formalised common sense and process knowledge, that no radically wrong decisions can be carried out. Of course, attempts can be made to learn all this from vast amounts of data by means of deep learning. But logic - and in particular the lightning-fast and error-free calculation with zeros and ones that forms the basis of computer science - offers a decisive advantage: one can use it to conduct long chains of reasoning precisely, guaranteed correctly and efficiently - "deep reasoning", which we humans have to learn onerously. This also allows conclusions to be drawn about rare phenomena that would otherwise be lost in noise and largely inaccessible for lack of sufficient data. So-called hybrid AI approaches therefore intertwine explicitly specified and learned knowledge, reasoning and learning.

It is clear that humanity will benefit enormously from third-wave AI systems. A good example can be found in climate research, a field that encompasses many disciplines. There are data and models for the atmosphere, the oceans and cloud systems. Economists are creating theories and models for climate change mitigation and adaptation. And there are biosphere models that deal with bacteria and decay processes of peat bogs and the Siberian permafrost. Do we really believe that a single person can put all these puzzle pieces together into a big, comprehensive picture, and make their findings comprehensible even to the general public? AI can offer crucial help with this important challenge. The same holds true for medicine. A third-wave AI system will be able to take into account all information about a patient, from demographic data, medication, laboratory tests and results of imaging procedures, to previous diagnoses and the specialist literature of all medical disciplines. The AI system then communicates its findings to the medical practitioner - and together the two develop a course of treatment that is individually adapted to the patient.

The third wave of AI offers enormous opportunities for Germany and Europe. The focus should clearly be on AI methods and applications that reliably expand human intelligence rather than trying to replace it, and that are as useful and accessible as possible to all parts of society. In this context, it is also important to ensure that AI research in the public sector - especially at universities and public research centres - can take place at the same level of excellence and ambition as in industry. Leadership in AI is too important to be left to private sector interests alone.

In this context, there needs to be a critical mass of expertise in the public sphere, also in Germany. In addition to the German Research Centre for Artificial Intelligence (DFKI), promising examples include the competence centres for AI designated by the Research Ministry in Berlin, Bonn/Dortmund, Dresden/Leipzig, Munich and Tübingen; the "units" of the European AI research network ELLIS in Berlin, Darmstadt, Freiburg, Heidelberg, Munich, Saarbrücken and Tübingen; and the partnership of universities, the Max Planck Society and industry that has emerged in the so-called Cyber Valley in Southern Germany. In Munich, more than 30 new AI professorships are currently being created, and recently, the Hessian Centre for AI has also bundled 22 new AI professorships and 22 founding members who already conduct AI research at its main location at TU Darmstadt.

The goal is to take a leading role in the third wave of AI by leveraging and intertwining the entire range of computer science, from hardware and robots to databases and software engineering to AI algorithms and exploiting synergies between natural and artificial intelligence. For example, following the idea of Josh Tennenbaum from MIT, one might assume that humans are born with a quasi-programmed, rough understanding of the world - in a way analogous to the game engines that exist for developing computer games. This "game engine in the head" provides the means for simulating the world and our interactions with it, and serves as the target of perception and the world model that guides our actions.

Similar to systems biology, we need systems AI: interactions of individual AI building blocks are mathematically and algorithmically correctly recorded, understood and used to obtain a valid overall view of complex AI systems.

Regional centres will play an essential role in the development of AI in Germany and Europe. In addition, there are plans to create a European lighthouse centre, a "CERN for AI", which will offer top researchers from all over the world a unique working environment and meeting place. Such a centre was first proposed in 2018 by the CLAIRE Initiative, which has meanwhile assembled the largest AI research network in the world, and is now featured in the European Union's plans. This would be a big deal. It is of crucial importance to resist the natural tendency of settling for a distributed infrastructure and working environment. Such a centre, which also has a symbolic character, cannot be distributed. It thrives on a critical mass of top scientists within the same physical working environment, on world-class support staff, and on unique technical infrastructure. Such a European institution could enable AI research in the public sector on par with market leaders in the private sector, such as Google, Apple or Facebook. Not only science, but also the economy would benefit from such an institution.

Europe actually has an excellent starting position - certainly also in comparison with the United States and China. What is missing is the political will to leverage this position in the best possible way to turn "AI made in Europe" into a major, trend-setting European success.

Holger Hoos is professor of machine learning at Leiden University in the Netherlands, co-founder of the CLAIRE initiative and vice president of the European AI research association, EurAI.

Kristian Kersting is professor of AI and machine learning at TU Darmstadt, co-director of the Hessian Centre for AI (hessian.AI) and a member of CLAIRE and ELLIS.