AI Challenges, Future of AI and Future Beyond AI: (notes from the IEEE Conference on AI CAI 2026 Granada, Spain, May 8–10, 2026)

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso
500 W. University
El Paso, TX 79968, USA
vladik@utep.edu, olgak@utep.edu

1 About the conference

1.1 A brief history of IEEE Conferences on AI

Artificial Intelligence (AI) is ubiquitous. Research in AI includes collaboration between many domains. Because of this, IEEE conferences on AI – that are being held every year since 2023 – are organized not by a single IEEE society, but by a collaboration between four societies: Computer Society (CS), Computational Intelligence Society (CIS), Signal Processing Society (SPS), and Systems, Man, and Cybernetics Society (SMCS).

The 2023 and 2025 conferences were held in the US Silicon Valley, the 2024 ones in Singapore. This year was the first time when the conference was held in Europe.

1.2 Who organized this year's conference

Each of the four conferences was organized under the leadership of one of the four IEEE societies. In particular, CAI 2026 was organized under the leadership of the Systems, Man, and Cybernetics (SMC) Society.

Two SMC leaders – Enrique Herrera-Viedma and Robert Kozma – served as this conference's co-chairs, as well as Carmen Nestares from Amazon.

1.3 Conference's sponsors, conference's participants

Amazon was not the only sponsor: more than 20 companies provided corporate sponsorship.

More than 450 people from all the continents (except from Antarctica) attended the conference, more than 30% of them young researchers. This percentage is a very good sign, because young researchers are a future of AI.

1.4 Keynote talks, workshop, etc.

The conference included three very interesting keynote talks:

- a talk "The Role of Algorithms in the Convergence of AI, Quantum, and High-Performance computing", by Juan Bernabé-Moreno from MIT Media Lab,
- a talk "AI Challenges in Mathematics of Investing", by Marcos Lopez de Prado from Cornell University and ADIA Lab, and
- a talk "Intelligent Computing for Cyber-Physical Systems", by G. Kumar Venayagamoorthy from Clemson University.

In addition, more than 300 regular talks were presented by authors from academic and industry; the acceptance rate was 52%. In addition to regular sessions, there were 12 workshops, several tutorials and interesting panels.

1.5 Some of the AI successes

Many of the talks described state-of-the-art successes of AI, successes that form the current state of AI.

For example, AI tools have been effective in the search for fusion-based plasma energy sources – the sources of the same energy that comes from our Sun and from other stars: LLM-based plasma models for promising designs like tokamaks outperformed traditional models in 13 out of 14 tasks, including an important task of predicting system's stability.

LLM-based models outperform more traditional models in many applications to materials, chemistry, weather and climate predictions.

1.6 AI challenges: the main focus of the conference, the main focus of these notes

Of course, the main emphasis was on the future of AI – what are the remaining challenges and what are possible ways to overcome these challenges. In this report, we provide a brief overview of these challenges and of the corresponding approaches.

Comment. It is impossible to overview 300+ papers in a short report – hopefully, other reports will cover what we missed.

2 Challenges and how to overcome them

What are the main challenges facing AI?

- Sometimes, AI is unable to solve the corresponding problem.
- Sometimes, AI is usually able to solve the problem, but in some cases, it provides wrong recommendation.
- Sometimes, AI provides good recommendations, but it uses too many resources – energy, water, etc. – to provide them.
- Finally, sometimes AI provides good recommendations, but people do not follow them.

Let us consider these challenges one by one.

3 First challenge: not enough computing power to solve some problems

3.1 Problem

The first challenge is that even with modern high-performance computers, for many important problems, there is not enough computing power to solve them.

3.2 Analysis of the problem

How can we further speed up computations?

One of the fundamental obstacles to increasing computing speed comes from physics, namely from relativity theory. Accrding to relativity theory, the speed of all processors is limited by the speed of light. This means, in particular, that in a simple laptop of 30 cm size, it takes at least 1 nanosecond for a signal to go from one side to another – and during this time, the usual 4 GHz processor already performs 4 operations.

To drastically speed up computations, it is therefore necessary to drastically decrease the size of all computer cells.

3.3 Need to take quantum effects into account

Computer cells are already at the size of several thousand molecules. If we decrease their size even further, we get to sizes comparable with the size of an individual molecule.

At these sizes, the usual Newton's laws of physics are no longer applicable. Such small objects obey special laws of the micro-world, known as quantum physics.

3.4 Advantages and challenges of quantum computing

One of the main features of quantum physics is that at these sizes, the behavior is no longer deterministic – it is *probabilistic*. Since we want the same – correct – computation result every time, this probabilistic character is a serious problem. So:

- on the one hand, we can have potentially faster computations, but
- on the other hand, we need to come up with new algorithms that would suppress the probabilistic character.

Such algorithm have indeed been invented, they form the basis of what is known as *quantum computing*.

3.5 How quantum computing can help AI

While the basic algorithms have been developed, the corresponding quantum hardware is still far from readiness. We do not yet have a universal quantum computer that would implement *all* known algorithms on the quantum level.

However, we already have quantum computers that implement *some* algorithms efficiently. So, to speed up AI-related computations, it is desirable to be able to use such quantum computers as much as possible – and, ideally, design new algorithms that would help us utilize quantum computers in AI.

3.6 How AI can help quantum computing

We have described how quantum computing can help AI – but AI can also help quantum computing, but helping to solve challenging tasks of coming up with new quantum algorithms and new designs of quantum computers.

3.7 AI and quantum computing need to collaborate more

We need to bring AI and quantum computing together. For this purpose, we need both AI and quantum computing specialists to learn each other's techniques and challenges.

Comment. Relation with quantum computing was the main focus of the keynote talk by Dr. Juan Bernabé-Moreno.

4 Second challenge: AI sometimes generates wrong recommendations

4.1 Fact

It is well known that AI tools often produce wrong answers.

- When ChatGPT was first launched, it produced a wrong answer 15% of the time.
- Now this portion is down to 5%.

This is better, but still not acceptable for many practical purposes.

For example, in emergency medicine, if we follow AI recommendations, then one out of 20 patients will get a wrong treatment and maybe even die. This is clearly not acceptable.

4.2 What can we do about it

We need to better educate people about the AI's errors. First of all, it is important to better educate people that AI can make wrong recommendations: many people still have over-trust in AI recommendations.

We need to teach people to detect AI errors. We also need to teach people to detect AI errors. For example:

- while it is important to teach students to use AI tools,
- it is also important to teach them to solve problems the old-fashioned way – with paper and pencil – so that they will be able to check for possible AI mistakes.

Similarly:

- in addition to teaching students how to use AI tools to program,
- we need to teach them to write programs without AI,

so that the students will be able to compare the results of the two programs and thus detect possible errors.

To decrease the effect of AI errors, we need to understand where they come from. To deal with the existence of erroneous AI recommendations, it is also important to understand where they come from. There are two main reasons for such recommendations:

- The first reason is that the objective function used to train an AI tool is often only an approximation to what we actually want. So, the recommendation which is perfect with respect to this objective function may not be that good from the viewpoint of our actual preferences.
- The second, more fundamental reason is that:
  - even when AI correctly detects the empirical dependence between the quantities,
  - it does not take into account that some of the dependencies are not causal.

  As a result, recommendations based on these dependencies sometimes make no sense.

4.3 First reason: wrong objective functions

Let us explain the first reason on four examples.

First example: AI tools never say "I don't know". One of the reasons why sometimes AI gives wrong answers is that in some situations, it does not have enough information to make a correct decision.

In such cases, why cannot the LLM honestly say "I don't know"? Because it was trained to maximize the number of correct answers. In situations when both "yes" and "no" are equally possible answers, randomly generating "yes" or "no" increases the average number of correct answers – and this is exactly what LLMs do.

We need to teach AI tools to say "I don't know" – and the only way to do it is to change the objective function that is used for training these tools.

Second example: AI tools are trained to please the user, not necessarily to tel the truth. AI tools whose objective is to please the users sometimes provide recommendations:

- that are bad for the users in the long term, but
- that help make the users happy.

This can be ignoring the bad consequences of the user's unhealthy habits, this can be providing biased (and often incorrect) information about the user's pet issues, etc.

Third example: AI has a potential to ruin the world's economy. We want economy to prosper. Empirically, capitalism – with some government control – seems to be the best scheme for increasing everyone's prosperity. From this viewpoint, it makes sense for each company to increase its profit. So far it mostly worked, but what if AI becomes in charge?

A natural way to increase the profit is to replace many humans by robots and/or AI tools. The company profits will increase, but if all the companies do that, the world's economy may collapse.

Is use of AI to help with research papers beneficial? At first glance yes, in reality no. Similar examples can be found in many areas in which AI is used.

For example, AI helps to write and edit research paper. A recent study has shown that, on average, papers helped by AI get more citations – and the number of citations is, at present, one of the main ways to gauge the quality of a research paper.

But the same study has shown that:

- if we use a more accurate criteria – e.g., if we ignore citations that point to the paper's faults,
- then, by this criterion, the impact of AI-helped papers on science is, on average, smaller than the impact of the papers that do not use AI.

So what can we do about it? There are many examples showing that AIs provide optimal decisions, but these decisions are optimal with respect to the goals that we give then. So, we need to be very careful in formulating the objective functions.

We need to clearly communicate our goals and constraints to the AI. We need to engage specialist in the application domain, specialists in ethics and economics to help better formulate out goals. And we need to monitor the AI results and detect when AI recommendations are misaligned with human values.

4.4 Second reason: correlation is not causation

The problem. The second reason is related to the known phrase: correlation is not causation. What current AI does is, crudely speaking, finding correlation between quantities.

For example, when the weather is good, ice cream sales go up, so is the number of people who go swimming – and thus, the number of drownings increases. An AI tool is perfectly capable to find the positive correlation between ice cream sales and drownings. Based on this dependence, if we ask AI how to decrease the number of drownings, it can recommend to decrease ice cream sales – which makes no sense.

How can we avoid this problem. To avoid such wrong recommendations, we need to supplement current AI techniques with techniques that detect causal relation. Such techniques exist: in a recent competition, most of the software tools correctly detected the causal structure. However, these techniques are still not perfect, they make mistakes in about 10% of the time.

There are also situations when causal analysis does not help: e.g., politics influence economy, and economy influences politics, both affect each others. In such cases, when there no tools that can help, we need humans to help.

Comment. The "correlation is not causation" problem was one of the main topic of the keynote talk by Dr. Marcos Lopez de Prado.

4.5 General recommendation

Automatic recommendations are sometimes wrong. In many such cases, humans can see that these solutions are wrong. So, a natural solution to this challenge is to avoid – at least for now – fully automated AI-based systems, and to have human experts monitor and check important AI's recommendations.

5 Third challenge: AI consumes too many resources – this is not sustainable

5.1 Problem

While AI techniques are very effective, they are very energy-hungry:

- while a human brain functions on about 20 Watt of energy,
- data centers already consume 5% of the world's energy production.

This proportion is expected to increase to 10% by 2030.

In addition to energy, LLMs use a growing amount of water for cooling their processors. All this is not sustainable.

5.2 What can we do about it?

What can we do about it? In a nutshell, there are two possible solutions:

- do not use AI where more resource-efficient tools are available, and/or
- make AI tools more resource-efficient.

5.3 First idea: sometimes, the use of AI is an overkill

Need to combine AI tools with more traditional techniques. A first natural idea is that for many problems, the use of AI and LLMs is an overkill: more traditional methods produce the results by using less energy.

In more complex problems, parts of the problem can be computed without AI with fewer resources. For example, computing products or logarithms on a traditional computer is usually much less resource-consuming that using a trained deep neural network to do it.

We need to be able to incorporate such traditional computations into AI computations.

Towards modular AI – and modular computations in general. Current AI techniques are similar to how some programs were written before software engineering: as one big program, for which any change requires redoing everything. Several decades ago, we learned that modular plug-and-play computing – in which we can replace a module without having to change everything – is a much more efficient way of computing.

A similar modular technique – known as Cellular Computing Networks (CCN) is being developed and used for AI. Some of these modules are pure AI, some can be more traditional computations. Such an architecture is also perfect for designing digital twins of a complex system by combining digital twins of its components.

5.4 Second idea: let us make AI tools more resource-efficient

Modular structure decreases resource consumption for some parts of the computation, but in this scheme, the AI part remains energy-hungry. So, for such modular computations schemes, the energy consumption decreases, but it still remain orders of magnitude higher than the energy efficiency of a human brain.

To make AI-related computations more energy efficient, let us recall the main difference between the signals in an artificial neural network and the signal in the biological neurons. While, in the first approximation, both neurons apply some activation function to the linear combination of inputs, a drastic difference is in how the input and output signals are represented:

- in a computer, we have a signal that continues for some time, while
- in the brain, signals are represented by very short spikes – and this saves a lot of energy.

It is therefore desirable to use spiking artificial neural networks instead of the current ones.

The idea of spiking artificial neural networks is not new, but it has not been much used in practice because there were no effective training algorithms for such networks. Now such training algorithms have been discovered. Networks using these algorithms require 10-100 times fewer resources than the usual artificial neural networks. so hopefully these networks will be, from now on, more widely used – and this will drastically decrease energy and water consumption.

Comments. Both ideas were among the main topics of the plenary talk by Dr. G. Kumar Venayagamoorthy.

6 Fourth challenge: AI provides good recommendations, but people do not follow them

There are two main reasons why sometimes, people do not follow good AI recommendations:

- First, many people have a natural distrust of new things – especially when these things are over-hyped. To solve this problem, we need to educate people about AI, we need to train them to effectively use AI tools.
- The second reason is more serious: people sometimes do not follow AI recommendations because, in contrast to human recommendations, the AI recommendations often do not come with any explanations. From this viewpoint, it is important to make AI more explainable.

7 How can researchers better face these challenges

7.1 Need for collaboration

Most AI challenges are too big to be handled by a single person, they need collaboration:

- collaboration with other AI researchers,
- collaboration with researchers from other disciplines,
- collaboration with industry.

Different team members have different ideas, and combinations of ideas are often useful,

Research collaborators are part of a person's support system; this support system may include other colleagues.

In these collaborations, it is important to take into account human factor: a collaboration is successful if we are surrounded by people who we enjoy to work with.

7.2 Interdisciplinary collaboration

To successfully use AI in solving a practical problem, there should be an interdisciplinary team including:

- computer scientists,
- specialists in the application domain, and
- if this problem involve humans – specialists in cognitive sciences, specialists in ethics, specialists in human-machine interaction.

Having people from other disciplines in a team also often helps to find possible new applications of the developed algorithms and products to new tasks and new domains.

Because of all this, students need to have experience in such collaborations.

7.3 Collaboration with industry

Industry is important for AI. Specifics of modern AI research is that it is largely done in industry, not in the academe:

- industry is where most openings for new PhDs are,
- indutry is where most of them end up working.

What AI researchers need from the industry. Traditionally, industry has always been more oriented on the results, especially on the immediate results. However, for research, there needs to be room for failure. Researchers need time and space to work on something which may not be of immediate use now.

We need thinkers, not hust doers. If a research project has a 100% success rate, this means that the tasks were not too challenging, a healthier approach is to have 70% success rate.

How AI researchers can enhance industry collaboration. To achieve the balance between short-term company goals and long-terms goals, researchers need:

- to educate managers about the importance of long-term research, and
- at the same time, to provide some focus on short-terms problems,

In other words, researchers needs to have a diverse portfolio of tasks – both low-risk and high-risk.

Researchers need to understand the planned character of industry work: a company has a strategic (long-term) plan and a morw short-term operating plan. Usually, anyone can make proposals to be included in this plan, with a clear indication of what resources are needed and what are the expected Key Performance Indicators (KPIs), and a proposed stopping point. Stopping point is important: the more you invest in an idea, the more difficult it is to stop.

Projects are periodically evaluated and, if needed, redesigned. At the end of each big project, there is an analysis: what did we learn? what could have been done better?

We need to better prepare students for industry jobs and industry collaboration. At present, the industry-related training of new hires largely starts when the person is hired by a company. However, to make new PhDs (and hew hires in general) more successful and more productive, it is desirable that the universities start better preparing students (in particular, doctoral students) for working in industry.

In addition to industry internships, universities need to more frequently invite industry representatives to talk to students, to make sure that students get experience in solving real-life industrial problems, problems based on real (or at least simulated) data. A good practice is to do what many new hires learn once they are hired: the experience of going from an idea to results in a short time – e.g., in a week.

Students also need to learn the economic side. Money is important. Engineers time is often more expensive than computation time, so if we can upload a problem to an LLM, this often saves money: simple solutions produced by LLMs are often good enough, and attempts to use humans to further improve them often do not pay off.

Computers also enable us to run computational experiments that check whether an idea will work, and LLMs can be used to evaluate the project;s results.

8 How to be better prepared for post-AI future

Universities prepare students for a lifelong career. At present, AI is one of the most promising directions, but in the future, something else may appear. We need to prepare students for this, we need more emphasis not only on the specific AI-related skills, but also on the general skills:

- Students need to have general *problem-solving skills*: how to start solving a problem, how to divide a complex problem into doable subproblems, etc.

- Students need to know the *fundamentals* such as mathematics and statistics that are needed to solve all kinds of problems. Time and again, it turned out that foundations are important. For example, a few years ago, some computer scientists seriously argues that computer science students do not need to know partial derivatives – and they were excluded from the ACM/IEEE recommendations. However, partial derivatives (and linear algebra) are a crucial part of backpropagation – the main algorithm for training neural networks.

- Finally, it is also important that students are able to continue working on a problem even when the original idea does not immediately work out. This *"grit"* is one of the main skills that students learn in graduate school.

In the fast-changing area of computing, specific knowledge often gets obsolete, so in the long term, it is not specific knowledge, but these general skills which are the most important.

9 Beyond science

The conference was held in a beautiful ancient-and-modern city of Granada, in one of the campuses of the 500-year-old University of Granada. The participants enjoyed not only a perfect organization of the conference and interesting talks from researchers from all over the world – including many from Granada, but also:

- great food at all the conference meals,

- an exciting flamenco show during the conference banquet,

- and – for many of the participants – interesting easy-to-walk-to tourist sites.

On behalf of all the conference attendees, we would like to thank the organizers for this great conference.