Systemic Environmental Risks of Artificial Intelligence

Description of expert assessment

("Leistungsbeschreibung")

As part of the project "Systemic Risks of Artificial Intelligence" we are issuing a call for proposals for an assessment of the systemic risks of artificial intelligence to the environment.

The understanding of systemic risks of AI seems to be in its beginnings. A generally recognized definition is still lacking. Threats to the functions of a system resulting from complex interactions between system elements are often seen as an important characteristic. Sectoral systems or industries (e.g., the financial system), the social system, or global systems are hold as relevant systems. There appear to be different types of systemic risks, which can differ, for example, in terms of their (un-)predictability during development or application of AI systems or whether damage can be attributed to a single actor.

To our understanding, systemic risks of AI can arise, for example, from the wide reach of AI providers or applications affecting large parts of society. They can also result from the propagation of potential harms from AI models in downstream applications or from the interaction between machines and/or humans. Furthermore, systemic risks may emerge from inappropriate stimuli or failures in collective behaviour leading to undesirable outcomes for society. However, the full range of systemic risks from AI may still be unknown, as they are largely unexplored and the technologies and applications of AI are developing rapidly.

Background

In recent years, the field and industry of artificial intelligence has accelerated dramatically. Being positioned as a disruptive general-purpose technology by influential entrepreneurs, policy-makers and academics, the sector has been able to acquire unprecedented amounts of venture capital. With newly announced investment programs and infrastructure construction,

efforts to push model performance ever further and realize widespread application in diverse areas show no sign of slowing down.

Whether for training models, expanding applications or fulfilling queries, the AI sector relies on vast supply chains, consumes considerable amounts of resources – energy, water, rare earth minerals and other materials – and produces waste. The socio-technical system of AI therefore carries a significant direct environmental footprint. Given that the global environmental and climate system is already under considerable stress and humanity is struggling to stay within earth system boundaries (ESBs), current and foreseeable AI resource needs and waste pose considerable environmental risks.

Aside from resource needs and waste production, some AI applications will also have further environmental impacts depending on area of deployment, e.g. agriculture. Beyond risks of harm, e.g. through accelerating fossil fuel exploitation, impacts of AI applications may also be positive, for example through efficiency gains, better climate modelling and monitoring, development of new materials, etc. Generally, there appears to be trade-offs between the environmentally significant inputs required by and potential benefits derived from AI applications. As AI is integrated into and reshapes a range of societal areas and practices, there likely are second-order environmental impacts, also in application context that do not have an apparent direct impact on the natural environment, e.g. digital advertising.

Environmental risks from resource extraction, hardware manufacturing, e-waste, as well as energy and water consumption of information and communication technologies (ICT) have been researched for several years. However, the extent of these risks, second-order effects and systemic dynamics specific to AI are under-researched, though there likely is some overlap with the broader ICT sector. In our project, we are interested in systemic environmental risks which emerge from the interaction between elements of the socio-technical system of AI: technical components, social actors, use cases and the environment. Such risks are likely not accounted for in design and training of AI models and applications and are usually very dependent on application contexts. Examples for systemic environmental risks are systemic externalities, rebound effects or path dependencies that make it harder for societies to adopt pro-environmental measures in the future. We expect further recognizable or reasonably foreseeable dynamics that can be described as systemic environmental risks. Particularly interesting are effects that contain feedback loops, and are non-linear, compounding or cascading.

For the report we are interested in AI models, systems and applications currently in use as well as reasonably foreseeable developments.

Guiding questions

1) Provide an overview of the current extent of environmental impacts (water, energy, materials, land use) of AI models, systems and relevant applications.

2) Detail sources, pathways, patterns and impacts of systemic environmental risks of AI and, where applicable, how they may spread. This includes but is not limited to rebound effects, path dependencies and systemic externalities. The inclusion and description of more phenomena of systemic environmental risks is highly welcome. Sources, pathways, patterns, impacts and pathways for spreading should be supported by evidence, for instance, in the form of literature reviews, case studies, or other means.

3) Describe relevant trade-offs present in AI systems and their applications, e.g. hardware or energy use and model performance (accuracy, response time, etc.). Are there areas where more resource-efficient models and systems are performing on a good-enough level (meaning higher resource use yields diminishing returns and no longer leads to meaningful improvements in model performance for that specific application)? Are there reasonably foreseeable developments that may reduce the resource needs of AI?

4) Compare (systemic) environmental risks from AI to those of the ICT sector at large. Are there differences in their characteristics or 'nature'? Are there best practices and governance measures that mediate the environmental impact of the ICT sector that are applicable to the AI sector as well? Or do systemic environmental risks from AI warrant specific or specialized governance measures?

5) How are the benefits derived from AI, including pro-environmental AI, and the environmental risks posed by the AI system distributed?

Context

The expert assessment is intended to contribute to a better understanding of the systemic risks of AI. It is part of the project "Systemic and Existential Risks of Artificial Intelligence", which is funded by the Federal Ministry of Education and Research (BMBF) (funding reference 01IS23075). The project is being carried out by the Institute for Technology Assessment and Systems Analysis (ITAS) at the Karlsruhe Institute of Technology (KIT). The project's research

aims to better identify, assess and avoid or mitigate systemic risks of AI and to derive insights for its governance.

As the assessment is to be produced in an interdisciplinary project context, the presentation of the expert assessment should be comprehensible to an interdisciplinary audience. It is expected that the assessment will be published by the authors (one or more co-authors) in a citable format in a timely manner after approval by ITAS.

ITAS is the point of contact for all scientific questions around the project and responsible for reviewing and approving the final assessment. Willingness to engage in intensive discussions and close cooperation with ITAS is a prerequisite.

Available financial means and deadlines

The maximum amount available for the expert assessment in the project is EUR 70,000.

- The deadline for submitting proposals is March 31, 2025.
- Work on the expert assessment is intended to begin on April 31, 2025.
- The expert assessment must be submitted to ITAS by September 30, 2025.

Notes on the preparation of the proposal

The proposal can be written in German or English. ITAS will review and scientifically evaluate the proposals and award the expert assessment. In order for ITAS to be able to evaluate the quality of the proposals, qualitative criteria must be considered when preparing the proposal. These criteria will be given equal weight in the evaluation:

- The proposal must demonstrate and document the particular expertise of the specific scientific personnel employed in the requested subject area in a detailed, clear, well-founded and transparent manner. In particular, the relevant scientific and research experience and/or other outstanding competencies (including acknowledgements and successes) in the subject area must be listed, both in terms of breadth and depth. Generally, this is to be demonstrated by presenting past projects with responsible accomplishment, activities relevant to the topic and (scientific) consulting services, as well as relevant publications.
- The overall quality of the content and form of the proposal will also be considered and evaluated. A clear structure is required. The planned effort and approach for preparing

the assessment must be clarified and justified in a detailed and comprehensible manner. Aspects listed in the call ought to be considered and addressed (as completely as possible).

- The description of the intended methodological approach for achieving the scientific expertise and work results relevant to the assessment will also be assessed. The chosen methodology and its particular suitability for the purpose of the assessment must be presented clearly and justified convincingly. The relation between the respective work packages, allocated time, and delivered content must also be transparent, clear, and justified.

- Lastly, the price of the respective proposals is also considered in the evaluation.

Please note the mandatory information that needs to be included in the proposal (see below). Please send your proposal as an electronic version to the e-mail address provided under 'Contact'. In our experience, detailed proposals often require revisions, e.g. with respect to formalities or calculations. If we shortlist your proposal after reviewing it, we will ask you to make the necessary revisions and then to send a signed written proposal to ITAS (P.O. Box 3640, 76021 Karlsruhe, Germany).

If you are awarded the expert assessment, a contract between ITAS and you will be drawn up and signed.

Contact

Lucas Staab

lucas.staab@kit.edu

Notes on mandatory information

In order to comply with the formal regulations of the KIT for proposals, please use the following wording for your proposal:

Proposal to the Karlsruhe Institute of Technology (KIT),

Institute for Technology Assessment and Systems Analysis (ITAS)

The following information must be included in your proposal:

- Name and exact address (no P.O. Box) of the proposing institution or person; for providers who work at a university or comparable public institution, but propose as a private individual, the private address is required.
- Function, title, first name and surname of the provider or authorised signatory (representing the institution, e.g. the chancellor in the case of universities/colleges)
- Exact title of the assessment
- If applicable, the person responsible for the assessment
- Date of the proposal
- Processing period: from ... to ...
- Date of submission of the assessment. Please note that the final version of the assessment will be delivered as an electronic version (doc or docx format), which also contains the original files of the tables and figures in the possible MS Office formats.
- Cost calculation including a separate VAT rate or a declaration that you are exempt from VAT. For personnel costs, the underlying time expenditure and estimated rates should be stated. The total price is treated as a fixed cost price.
- The proposal and further documents can be submitted electronically as PDFs.
- A short CV of the persons working on the project and, if applicable, a short introduction of the providing institution should be included as an attachment.

We are looking forward to your proposal!