Guidelines on the Use of Artificial Intelligence in Research

Preamble

Artificial intelligence (AI) can be used today in many fields to provide context-dependent support. AI encompasses applications in which machines perform tasks that approach, match, or exceed human cognitive capabilities. In the context of these guidelines, we use the term specifically to refer to generative AI models, including AI agents, as well as machine learning, including deep neural networks.

Ethical AI refers to AI systems that operate in compliance with the law, are technically robust, environmentally sound, reliable, and free from bias; that avoid discrimination; and that allow for human oversight. Such systems are not the norm. Therefore, it is important for us to adopt a thoughtful, nuanced, and responsible approach to AI, as it presents both opportunities and risks. We take care to protect sensitive data and information and are aware of potential bias and misinformation (hallucinations). For us, the foundation for working with AI is formed not only by the EU AI Act but also by the Rules to Secure Good Research Practice and Avoid Academic Misconduct at the University of Rostock, as well as data protection [DE], research security [DE], and copyright law. As researchers, we take responsibility for adhering to these rules and support one another in doing so. As part of our sustainability goals, we ensure that our use of AI is resource-efficient.

Scientific Work

The personal contribution to scientific work is being transformed by the use of AI. The competencies required for scientific work are discipline-specific, but in every case include an understanding and evaluation of scientific methods as well as the “critical mind” with which ideas are put to paper. We do not replace these competencies with AI.

Depending on the discipline, AI tools can be used in data collection and analysis itself and/or in academic writing, the writing of expert reports, and the preparation of grant proposals.

Data Collection and Analysis

When using AI in research, we critically examine the algorithmic fairness of methods, data, and results.

We are aware of potential biases - for example, in the training data and, consequently, in the results - and understand that this awareness is a fundamental prerequisite for the responsible use of AI. We handle data with care and keep in mind that the data entered is usually reused by AI tools for training and personalization. This means, for example, that we do not upload any personal or confidential data to AI systems unless it is ensured that the data will not be stored, reused, or processed for other purposes there. When in doubt, we use local models that comply with data protection regulations or those with zero data retention. We adhere to the terms and conditions of the software providers.

We are aware that AI can infringe on copyrights during data mining and content creation. If AI-generated results are adopted without critical review, this can result in copyright infringement and, in some cases, plagiarism. Therefore, we take our responsibility as authors for the content of the publication seriously. We are supported by the Data Protection Officer [DE] and the Ethics Committee [DE] in addressing questions regarding the use of AI in research.

We understand that AI technologies operate on a probability-based model. The results of AI-generated responses are not reproducible.

We take responsibility for our research data and observe ethical, data protection, and confidentiality considerations [DE] in our research data management.

Scientific Writing

In academic writing, we, as authors, take responsibility for the text. We disclose the use of AI in the creation of text, figures, and tables, as well as in the editing of texts or similar activities. Transparency is important to us. We cite sources in accordance with discipline-specific rules; when using AI, we cite it as a source with the standard citation information. We prefer to cite the original sources rather than the AI itself. When using AI as a method, we describe this in the relevant section of the text and, if applicable, also specify the prompt used.

Many publishers already have guidelines for the use of AI and for citing its use. We adhere to these guidelines.

Doctoral dissertations and habilitation theses represent a unique aspect of scholarly work: the goal here is to demonstrate independent scholarly work. AI-generated texts do not constitute independent work, and we do not use them without careful consideration.

Writing Expert Reports

Writing expert reviews is a core task for us as scholars, one that requires our subject-matter expertise. We can use AI as a supportive tool, for example, to convert keywords we have written ourselves into continuous text or to search for further information. We do not delegate the evaluation itself to AI. Furthermore, we maintain confidentiality. We do not upload grant applications or manuscripts to AI tools unless it is ensured that the information will be handled confidentially and that the data will not be stored and/or used for other purposes. As reviewers, we assume responsibility for the review.

Submitting Applications for Third-Party Funding

External funding agencies publish guidelines on the use of AI in writing and reviewing grant proposals. We adhere to these guidelines in their currently valid form.

We may use AI to generate or refine ideas, provided this is done in compliance with data protection regulations and in accordance with the Rules to Secure Good Research Practice and Avoid Academic Misconduct. We also take research security into account: We are aware that the data entered is reused by AI tools. This can lead to difficulties if it results in a violation of export control [DE] regulations.

Furthermore, AI can be used to refine texts, provided the same rules are followed and the text is appropriately labeled (see Academic Writing).

GLOSSARY

Bias refers to cognitive distortions in perception, such as prejudices, stereotypes, and other errors in reasoning.

Generative AI methods are AI applications that generate content in the form of text, images, audio, video, software code, or datasets in response to an input (prompt). They are typically trained on large datasets. They are probabilistic, which not only produces variations in the results but can also lead to hallucinations (incorrect results).

Hallucinations (also referred to as misinformation in the context of artificial intelligence) are convincingly phrased but factually incorrect or entirely fabricated statements generated by language models.

AI agents are autonomous programs that not only answer questions but also perform complex tasks on their own, such as processing emails or writing autonomously.

Machine learning (ML) encompasses statistical methods that can autonomously recognize and, if necessary, learn patterns in data. This pattern-recognition capability allows for the analysis of relationships and, where appropriate, the derivation of predictions without explicit, hard-coded instructions.

(Artificial) **neural networks** are a subfield of machine learning. Neural networks consist of tightly interconnected artificial “neurons,” which today are complex, mostly nonlinear, statistical models. As a result, results—particularly in deep neural networks—are not fully reproducible.