# Guideline on the Use of Artificial Intelligence in the Review Process

Disclaimer: The English translation of this document is provided for informational purposes. In the event of a discrepancy between the English and the German versions, the German text takes precedence.

DFG

**Preamble**

The review of funding proposals is an integral element in determining scientific quality in connection with the DFG's research funding activity. (Generative) artificial intelligence (hereinafter referred to as "AI systems") is now widely available and can be used without specific technical expertise. In view of the need to protect the legitimate legal interests of applicants, maintain the confidentiality of the review process, meet general transparency requirements and ensure the research integrity of DFG funding programmes as a whole, the use of AI systems is permissible only in accordance with the principles set out below.

This usage and transparency guideline aims to establish a framework for reviewers using AI systems in the review process so as to ensure that AI systems are used in a way that is responsible, transparent and – with regard to legislation such as the AI Act (EU) 2024/1689, the General Data Protection Regulation (GDPR) and the German Copyright Act (UrhG) – legally compliant. Responsibility for the content of the review remains with the reviewer.

**Purpose of use**

Human analysis and the expert assessment of the proposal are core elements of the review process. The use of AI systems in the preparation of reviews is therefore permitted only in a supporting capacity. All substantive judgements regarding the funding recommendation rest solely with the reviewer.

*For example, an AI system may be used to identify relevant literature relating to the proposal under review, to turn reviewers' own keywords or notes – from which the intended direction of the review is already clearly apparent – into a fully formulated text, or to provide suggestions for structural or linguistic revision of the reviewer's own ideas.*

*However, it is not permissible to have substantial parts of the review generated automatically and without reflection, i.e. without independent expert assessment.*

# Guiding principles for the use of AI

## 1. Confidentiality

AI systems may be used only where care is taken to ensure that confidential content of funding proposals is not processed without oversight, stored, or used for purposes other than the conduct of the review process. The supportive use of AI in drafting a review must not result in the disclosure of confidential proposal content. This information is entrusted to the DFG exclusively for the purpose of review, evaluation and decision-making.

*In particular in the case of cloud-based AI systems, subsequent processing for training purposes may occur when data is entered. This is not covered by the original purpose for which the information was provided by the applicants and is therefore not permitted.*

*For this reason, reviewers may input proposal content into AI systems only if care is taken to ensure that the provider of the AI system does not store the proposal content permanently or beyond the specific purpose of use.*

*One of the following requirements must therefore be met:*

- *The AI system must be installed on the reviewer's own device and processing must take place exclusively on that device, without cloud computing or the transfer of processing results to the cloud.*
- *The AI system must be hosted by a trustworthy institution (e.g. a university or public research institution).*
- *Cloud-based non-commercial or commercial AI systems (e.g. ChatGPT, Claude, Gemini) may be used only where a contractual licence agreement guarantees data security and excludes use of the input data by the provider (e.g. for training purposes), and where these conditions are actively set by means of the appropriate security settings (e.g. "no logging"; "no storage").*
- *The recommendation is that AI systems should be used which are based on transparently documented and publicly available AI models with publicly accessible training data.*

## 2. Transparency

If an AI system is used in preparing the review, this must be disclosed, specifying the purposes of use.

*The use of the AI system can be described in a few explanatory sentences. Specific marking of the review passages in question is not required. "Disclosure" here means stating which of the permissible purposes listed above the AI system was used for and to what extent.*

*AI tools that do not affect the substance of the review (e.g. grammar, style or spelling checks, translation tools) need not be disclosed.*

### 3. Quality assurance

Uncritical adoption of AI-generated text is not permitted. AI-generated text must always be critically assessed, adapted and checked for factual accuracy, currency and potential bias structures.

> *Since generative AI models are based on probabilistic methods, they may produce factually incorrect, unbalanced or biased content. The models are trained on publicly available data and use probabilistic calculations to reflect social, cultural and also academic discourses without taking an academically substantiated position of their own. As a result, they perpetuate bias contained in the training data, including gender and diversity bias, for example. With regard to the processes of knowledge acquisition by research, they may evaluate concepts that build on established approaches more positively than novel research approaches that could lead to paradigm shifts (status quo bias). Overall, generative AI may produce inaccurate or distorted content. Without critical human control, the scientific quality of the review and the integrity of the DFG review system are at risk. Seemingly coherent AI-generated text may also give rise to automation bias.*

### 4. Responsibility

Full responsibility for the content of the review lies with the person who prepared it. This applies even where parts of the text were generated with the assistance of AI systems. Responsible expert review requires subject expertise and ethical judgement, which by their very nature AI systems do not possess. Responsibility must therefore not be transferred to an AI system; it remains with the reviewer.

> *These principles are specified in the DFG Code of Conduct [Guidelines for Safeguarding Good Research Practice](). Guideline 16 identifies compliance with the confidential handling of third-party content submitted for review or evaluation as a key element of the process of forming a scientific judgement. Researchers submit unpublished scientific ideas, research data, illustrations and methodological details that are often the result of years of work and therefore particularly worthy of protection. Protecting such content is essential in order to safeguard the rights of authors, maintain trust in a science-driven review process and prevent scientific misconduct.*
>
> *Guideline 16 identifies neutrality of review processes as an important basis for credibility, fairness and quality in scientific decision-making processes. This is the indispensable foundation for a fair and transparent evaluation process in research. Only when reviewers present their views independently, objectively and without undue influence can academia and society have confidence in the outcomes of such processes.*

*Under § 9 No. 2 of the [Rules of Procedure for Dealing with Scientific Misconduct](#) (VerfOwF), reviewers may potentially be in breach of confidentiality and thereby commit scientific misconduct if they disclose scientific content entrusted to them for review to unauthorised third parties. This may occur where content is entered into large language models (LLMs), if the content is used as training data.*