WIAS guideline on the use of generative AI

Preamble

The use of *generative artificial intelligence* **(GenAI)** has largely become established in society and science. There is a consensus that this disruptive technology has great positive potential for teaching and research, which should be realised, managed, and accompanied at WIAS. The goal of scientific GenAI use must be to gain knowledge that strengthens society's trust in scientific methods and their results while preserving the freedom of research and teaching. The use of GenAI at WIAS is primarily determined by:

- EU Artificial Intelligence Act [EU AI Act],
- EU *General Data Protection Regulation* [GDPR],
- *Guidelines for Good Scientific Practise* by DFG and Forschungsverbund Berlin e.V. **(FVB)** [GWP],
- *Living guidelines on the responsible use of generative AI in research* [EU Guideline],
- *Recommendations for safeguarding good research practice when using artificial intelligence* [Leibniz Guideline],
- guidelines of journals, partners, and universities.

This guideline is intended for WIAS employees in research, administration, and IT, as well as for WIAS guests who use WIAS GenAI systems. It primarily regulates the official use of the so-called *large language models* **(LLM)**. The core principles of research with and on AI are [EU Guideline]:

- **Reliability:** Ensure research quality through verifiability and reproducibility and be mindful of bias.
- **Integrity:** Research should be transparent, fair and impartial; the use of generative AI should be disclosed.
- **Respect:** Respect for people, society, the environment, sensitive data (e.g. personal data, trade secrets, expert opinions, applications, copyrighted content), privacy, intellectual property and fairness.
- **Accountability:** Researchers remain responsible for all AI-generated results and their consequences; generative AI is used under human supervision and control.

Prohibited AI practices within the scope of the [EU AI Act] are also excluded from research at WIAS; research and development at WIAS is not directed towards such purposes.

With this guideline and the accompanying processes, WIAS complies with the obligations under Articles 4 and 11 of the [EU AI Act] to ensure an appropriate level of AI competence and to provide technical documentation for the use of the AI systems offered. Overall responsibility lies with the institute's management, and implementation is carried out by the roles specified in this guideline (including data protection coordination, Ombudsperson, RSE, IT). Further training on AI topics will be offered in the future, *e.g.* as part of the RSE seminar at WIAS, there are planned events to ensure AI competence.

Tier system (TIER 0,1,2)

Within the scope of the [EU AI Act], *scientific research* is generally associated with *low risk*. WIAS uses a TIER tier system for further practical assessment and limitation of risks when using LLMs. It distinguishes between the type and purpose of the LLM, but does not replace the official high-risk/low-risk classification (and prohibited practices) of the [EU AI Act]. The main features of the TIER levels are shown in Table 1. Depending on the complexity of the research questions, researchers can use TIER 0, TIER 1 or TIER 2 LLM. The development, adaptation and evaluation of proprietary AI models for research is expressly permitted within the framework of the TIER levels. In doing so, the relevant legal provisions [GDPR, EU AI Act] and the relevant guidelines [GWP, EU Guideline, Leibniz Guideline] need to be complied with. The GDPR and EU AI Act define certain exceptions and privileges for research (e.g. Art. 89 GDPR or Recital 25 EU AI Act).

The use of LLMs in administration is subject to the legal requirements of [GDPR, EU AI Act]. In particular, processes in human resources and financial management, third-party funding and contract management, and internal organisation are subject to a special *duty of care*. For administrative purposes, therefore, only TIER 0 and, in cases of clearly non-sensitive[1] data should generally be used for TIER 1 LLM. AI systems may not be used for the automated

[1] Clearly non-sensitive data does not contain any personal, copyrighted or confidential information.

TIER	Risk	Typical models / use	Requirements
0	negligible	local LLM without external data transfer; basic tools for text and code support	automatic access; compliance with guideline (**black**) and [GWP, GDPR]
1	low	curated, external LLM with standard functions (chat, text, code) for research and administration without sensitive data	acknowledgement & compliance with guideline (**black**) and [GWP, GDPR]; restriction to non-sensitive data
2	medium	powerful external LLM, specialised APIs, complex agentic workflows; Research *with* and *on* models	application & approval; compliance with TIER 2 guideline (**blue** + **black**) and [GWP, GDPR], extended requirements/expertise

Table 1: Tier system for AI use cases at WIAS

evaluation of applicants or employees.

For TIER 1 and TIER 2, the token and cost limits of the external LLM providers must also be observed: the models must be used in an energy- and cost-saving manner, differentiated according to their intended use. In TIER 0 and TIER 1, use is exclusively from the WIAS intranet/VPN via

`https://webui.wias-berlin.de`.

In TIER 2, access to external LLMs may be direct and therefore not necessarily from the WIAS intranet/VPN. The approval process for TIER 2 is carried out by means of an **application** (see Table 1).

Guideline

Researchers at WIAS can use generative artificial intelligence like other tools to help with research, teaching, software development, and administration, as long as they follow the rules of the [GWP], [GDPR], universities, and journals. Responsibility for the accuracy and usability of the content (input and output) always remains with the AI user. All WIAS employees must observe the applicable legal regulations when using AI, e.g. with regard to data protection, copyright and patent law.
The next section highlights key aspects of GenAI use.

Rules of conduct and documentation in research

1. The guidelines for good scientific practice of the Forschungsverbund Berlin e.V. and the DFG [GWP] form the basis of all scientific work at WIAS. GenAI must not replace original scientific work, but only support it. Outputs must be checked for accuracy (hallucinations) and bias.
2. The use of GenAI for scientific work (publications, expert opinions, reports, teaching materials, etc.) must be documented in such a way that the type, scope and purpose of the support or the individual scientific achievement can be traced.
3. Authorship, responsibility and liability remain with the respective users and, where applicable, co-authors. GenAI systems cannot be an author.

TIER 2: In order to make the risks associated with use traceable, a research data management plan (RDMP) appropriate to the complexity of the data must be drawn up. The RDMP is adapted if the data profile changes. In simple cases, the application serves as the RDMP. Particularly in TIER 2, a high level of AI expertise is required, and users are required to keep their AI expertise up to date.

Data protection When using AI systems, the same data protection principles and requirements apply without restriction as for all other applications and processes. The following aspects in particular must be observed to ensure that AI systems are used in compliance with data protection regulations:

4. The user is responsible for the legality of the inputs into and the utilisation of outputs from LLM. External LLM (from TIER 1) can store inputs and outputs for some time and monitor them automatically to check for legal violations. Inputs sent via the API to external LLMs are, according to current information, not used for training. However, feedback such as thumbs up or thumbs down may be regulated separately and may still be used for training under certain circumstances.

   Details for the curated TIER 1 LLMs can be found in the terms of use of the respective providers (e.g. Mistral or OpenAI).
5. The processing of personal data in the sense of the [GDPR] or sensitive data requires a legality check with regard to applicable laws and internal data protection regulations. This must be done in consultation with the data protection coordinator and the FVB data protection officer.
6. If personal data may be processed in research projects, it must be anonymised or pseudonymised in an appropriate manner depending on the type and nature of the data. In addition, the obligations imposed as a result of the approval process, for example with regard to data minimisation, data security, documentation, deletion periods and the rights of data subjects, must be implemented.

7. Data protection incidents: If a breach of personal data protection has occurred, this must be reported immediately to the data protection coordinator and steps must be taken to ensure that protection is restored.
8. If you have any questions or concerns, please contact the data protection coordinator or the FVB data protection officer for assistance.

TIER 2: Depending on the use case, users of TIER 2 projects have a special obligation of due diligence (e.g. curation of LLM models used, overview of shared/generated data, configuration of LLM services (opt-out training, data storage)). When using external providers in countries outside the EU, the requirements of the [GDPR] for third-country transfers are reviewed by data protection coordination.

Software development and copyright

9. GenAI can be used to assist in the development of software (e.g. code, shell scripts, websites). Responsibility for the functionality, correctness, security, copyright and licence compliance of the resulting software lies with the developers/AI users.
10. Ensure that no confidential source code or proprietary algorithms are disclosed in TIER 1/2 unless expressly authorised to do so. For open source projects, the respective licence terms must be observed. Note: *GenAI can contain/reproduce copyrighted content. Many commercial and open LLMs have no transparent documentation of the training data used, which makes it significantly more difficult to verify copyright.*

TIER 2: The use of AI in software development should be documented in a comprehensible manner, for example through meaningful commit texts. Users should check the compatibility of the components used with licence and copyright requirements. Copyright arises exclusively from human creative work; codes generated purely by AI are not protected by copyright. Additional technical protective measures are recommended for agentic systems that automatically generate and execute code or control external tools (e.g. sandbox environments, restriction of critical actions, logging).

Teaching and publications

11. In teaching, GenAI can be used as a support tool by both teachers and learners. Exam and study regulations as well as the respective provisions of the universities must be followed. Deceptive practices (e.g. submitting AI-generated texts as your own work without labelling) are not allowed.
12. For scientific publications, the use of GenAI must be made transparent if required by journals, publishers or funding organisations. The applicable guidelines (e.g. on the use of AI in manuscripts) must be adhered to.

TIER 2: When used in publications (e.g. extensive AI-supported data analyses or complex agentic workflows), it must be ensured that the results remain traceable, reproducible and independently verifiable. This includes, in particular, the documentation of the models, configurations and data processing steps used.

Concluding remarks

This guideline shall be reviewed regularly and amended as necessary. In particular, European and national regulation of AI, new systems from external providers and technical developments (e.g. new model classes, agentic systems) require ongoing further development of the institute's internal framework.

When developing proprietary AI models and AI systems, unless these are used exclusively for research purposes within the meaning of Article 2(6) of the AI Regulation, the provider obligations under the AI Regulation must also be reviewed and observed. Changes to the TIER system, the approved models or the roles and responsibilities (e.g. Ombudsperson, data protection coordination, IT) are recorded in the corresponding TIER 1/TIER 2 guideline and in accompanying process documents. In cases of doubt, this guideline refers to these supplementary documents.

The Ombudsperson, data protection coordinator, research software engineer (RSE) and IT department are available as contact persons for questions regarding the interpretation of this guideline, specific use cases or planned TIER 2 projects.

Authors: C. Merdon[2], D. Peschka[3], L. Prieto[4]

[2] Data Protection Coordinator at WIAS
[3] Ombudsperson at WIAS
[4] Research Software Engineer (RSE) at WIAS

References

[EU AI Act] EU KI-Gesetz / Artificial Intelligence Act (Verordnung (EU) 2024/1689). https://artificialintelligenceact.eu/de/ und https://eur-lex.europa.eu/eli/reg/2024/1689/oj

[GDPR] *Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr (Datenschutz-Grundverordnung).* https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[GWP] Deutsche Forschungsgemeinschaft (DFG): *Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“*. https://www.dfg.de/gwp. Forschungsverbund Berlin e.V. (FVB): *„Verfahrensordnung bei Verdacht auf wissenschaftliches Fehlverhalten“* https://www.fv-berlin.de/ueber-uns/hinweise-kritik/gute-wissenschaftliche-praxis

[EU Guideline] European Commission: *Living Guidelines on the Responsible Use of Generative AI in Research*. https://research-and-innovation.ec.europa.eu/document/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en.

[Leibniz Guideline] Leibniz-Gemeinschaft: *„Empfehlung zur Sicherung der guten wissenschaftlichen Praxis beim Umgang mit Künstlicher Intelligenz“*. DOI: https://dx.doi.org/10.5281/zenodo.14420893.