WIFO AI Guidelines on the Proper Use of Generative Models of Artificial Intelligence

Michael Peneder
Fabian Unterlass

WIFO AI Guidelines on the Proper Use of Generative Models of Artificial Intelligence

Michael Peneder, Fabian Unterlass

Austrian Institute of Economic Research

Internal review: Christine Mayrhuber

Abstract

Artificial intelligence presents new challenges for both organisations and individuals in terms of how to navigate and manage its use. As the guidelines developed at WIFO may be of interest to other research organisations facing similar challenges, we have decided to make them publicly available. WIFO's AI Guidelines provide a comprehensive framework for the ethical and effective use of generative AI in research and administrative support. The guidelines stress the importance of continuous education, transparency, and scientific integrity, forbidding AI in personnel assessments and involving employees in AI governance. Key components include robust technical infrastructure, an AI working group, and internal communication channels. The guidelines highlight personal accountability, human oversight, and strict data protection, ensuring compliance with European laws and fostering responsible AI integration.

WIFO AI Guideline
on the Proper Use of Generative Models
of Artificial Intelligence

Michael Peneder, Fabian Unterlass

Artificial intelligence presents new challenges for both organisations and individuals in terms of how to navigate and manage its use. As the guidelines developed at WIFO may be of interest to other research organisations facing similar challenges, we have decided to make them publicly available. WIFO's AI Guidelines provide a comprehensive framework for the ethical and effective use of generative AI in research and administrative support. The guidelines stress the importance of continuous education, transparency, and scientific integrity, forbidding AI in personnel assessments and involving employees in AI governance. Key components include robust technical infrastructure, an AI working group, and internal communication channels. The guidelines highlight personal accountability, human oversight, and strict data protection, ensuring compliance with European laws and fostering responsible AI integration.

1. Introduction

The proper use of artificial intelligence (AI) offers great opportunities to increase efficiency, effectiveness, and creativity—not only in research, but also across all supporting services, such as research administration, support, and public communication. However, the rapid pace of technological change demands a high degree of attention, flexibility, and a willingness to engage in continuous learning.

The following guidelines summarize WIFO's current principles and practices concerning the use of AI within the institute. They are part of an ongoing process and are regularly reviewed and adapted to reflect new developments. The focus is on generative AI models, which are distinguished by their ability to learn patterns, structures, or properties from existing data and to generate new, original content based on them.[1]

The WIFO Management Board has explicitly committed not to use AI models in administrative processes, such as personnel information systems, for evaluating or assessing employees. Furthermore, WIFO ensures workforce participation by either including a works council member in

---

[1] The definition of the term "generative AI models" used here is based on a query to ChatGPT as well as further research (Google, Wikipedia) conducted on March 14, 2025.

the WIFO AI Working Group (Section 2) or involving the works council in relevant decisions that may affect employee rights as defined in § 96 (1) (3) *ArbVG* (Austrian Labour Constitution Act).

2. Framework conditions

WIFO aims to create optimal framework conditions regarding resources and processes for the use of AI across all areas of work. This includes (i) ensuring a reliable, secure, and high-performance technical infrastructure, (ii) providing employees with sufficient training and continuing education opportunities, (iii) establishing transparent organisational structures, processes, and responsibilities, and (iv) fostering a dynamic environment for knowledge exchange and the development of new ideas:

- The WIFO **AI Working Group** is the primary point of contact for both management and employees regarding requests, suggestions, and the exchange of experiences with new AI tools. Before new generative models are implemented, the group evaluates them for practical utility and safety, especially with regard to data protection. The working group acts on behalf of the management and, in close coordination with the Chief Information Officer (CIO), provides recommendations on the adoption of AI tools. In areas that are less critical, involve low costs, and carry no apparent operational risks, such as those related to data protection or copyright, the group is authorized to make decisions independently or in consultation with the CIO.

- **Technical infrastructure**: In alignment with European guidelines (EU 2023) and to ensure the protection and confidentiality of input data, generative AI models should be hosted locally or operated directly by WIFO, wherever this is technically and economically practicable. If this is not possible, all applicable legal requirements and WIFO's internal policies for the use of AI models – with particular emphasis on data protection, personal rights, and IT security (see Section 4) – must be applied with particular rigour. These rules are especially stringent for external AI systems, but naturally also extend to all self-hosted models.

- The **WIFO AI-Park** is an internal digital platform for institute-wide use of artificial intelligence. It is organized by application areas, such as research and coding, and provides a structured overview that includes access information, practical tips, and use cases for all tools. The AI working group manages the AI-Park and continuously reviews suggestions for expanding its offerings. Input from employees is welcome and forms an important basis for the further development of the platform. To balance the tension between rapid innovation and due diligence, the AI working group clearly distinguishes between tools that have already been tested and those that are still under evaluation. The latter are marked with appropriate warnings but can still be used, promoting feedback to improve the evaluation process and helping to develop the AI-Park in line with user needs.

- **Education and training**: WIFO provides a comprehensive range of AI training and education for its employees. In accordance with the *European Artificial Intelligence Act* (EU 2024), these programs address not only technical and practical competencies, but also essential legal topics such as data protection, privacy, and intellectual property.

Furthermore, they convey a fundamental understanding of how AI systems operate. The aim is to empower employees to apply AI-generated results both independently and with a critical mindset, ensuring responsible and informed use in their daily work.

- ▪ **WifoGoesAI:** This initiative provides all interested employees at WIFO with an independent communication channel through MS Teams, which is moderated by the AI working group when necessary. The informal exchange of information and experiences is intended to encourage and support the use of generative AI systems at WIFO. This channel also serves to continuously adapt tools and guidelines for the use of generative AI to new developments, to identify training needs, and to prevent potential misuse of AI tools. Employees are therefore encouraged to participate as actively as possible in this exchange.

3. Scientific integrity

In addition to the requirements for technical infrastructure and personal competencies in working with AI, scientific integrity is a key prerequisite for ensuring research quality, trust, and a positive societal impact. The use of AI, like all other research activities, must comply with all applicable national, European, and international legal regulations. The same standards and rules apply as those outlined, for example, in the European Code of Conduct for Research Integrity (ALLEA, 2023) and in WIFO's guidelines on scientific integrity. Any form of misconduct, such as fabrication, falsification, or plagiarism of data and results, as well as violations of personal rights, copyright, or data protection,[2] is strictly prohibited. These principles are relevant to all areas of work at WIFO. Moreover, scientific integrity imposes the following specific requirements on research:

- • **Transparency** regarding clients, financial contributions, and conflicts of interest
- • Proper **citation** of all relevant sources for ideas, texts, data, and other materials
- • Thorough **review** of results, taking into account critical assumptions, robustness, and the limitations of the methods used
- • **Traceability** of results by retaining all necessary documentation (such as data, methods, questionnaires, and sampling plans) for at least 10 years

4. Specific new challenges

The wide range of new opportunities, together with the unprecedented and rapidly growing capabilities of generative AI models, presents additional challenges (EU 2023, DFG 2023). The following principles therefore complement and clarify WIFO's guidelines on scientific integrity in relation to the use of generative AI. These principles apply both within the WIFO AI-Park and, whenever the use of AI tools is connected to WIFO activities, also outside of it.

---

[2] See, e.g., Birnbauer (2025).

Personal responsibility

The legally compliant and rule-abiding use of generative AI lies within the personal responsibility of the individual actively employing it or further applying its results. This responsibility also includes a willingness to engage in continuous learning and to improve one's skills in handling generative AI models, for example through training sessions or participation in user forums. In line with this personal responsibility, every employee who uses generative AI models is accountable for the integrity of the content they use by:

- carefully reviewing AI-generated results,
- ensuring that no AI-generated fabrications, falsifications, or plagiarized content are used or disseminated,
- maintaining a critical approach when interpreting the results, and
- being aware of the limitations of the tools, such as potential hallucinations, biases, inaccuracies, and lack of reproducibility.

In this context, organizations such as the *Leibniz Gesellschaft* (2024, p. 3) emphasize that AI systems are merely supportive tools. Their results, *per se,* lack scientific validity and cannot substitute for authentic processes of gaining knowledge or making decisions. Therefore, any new AI-generated content must always be carefully reviewed for accuracy by the responsible individual. If justified doubts about its accuracy remain, the use of AI tools should either be limited to initial idea generation or brainstorming, followed by comprehensive and independent research, or avoided altogether.

This personal responsibility establishes the principle of **Human in the Loop** for all applications of generative AI models, meaning the primacy of human oversight and accountability. According to this principle, a responsible individual must be involved at both the beginning and the end of every application.[3] For example, this means that there must not be any automated publication of AI-generated content without prior review by an employee responsible for that content.

Transparency

The **authorship** of content generated with the help of generative AI models remains with the human users of the AI.[4] However, when generative AI is used in the research process, it is necessary to clearly indicate in any publication which tools were used, when they were used, and for what purpose. The rules for citing AI-generated content are still under development internationally and are updated regularly. These rules can differ significantly between research societies and academic publishers. For publications, the specific guidelines of the publisher, the editors of research results, or the developers of the AI tools used must always be followed. The following rules and guidelines provide general orientation for work at WIFO:

---

[3] See, e.g., APA (2023) or ORF (2025).

[4] For example, copyright protection also requires that "these are (i) original and (ii) intellectual creations of a human being" (Birnbauer, 2025, Reg. 6, Chap. 1.2, p. 1).

- If the data is not publicly accessible, it should be treated as **personal communication** and cited either in footnotes or directly in the text (e.g., ChatGPT, April 6, 2025).[5] Inclusion in the bibliography is not necessary.

- The use of purely technical tools that do not generate original content may be exempt from this requirement.[6] The responsibility for assessing whether AI-assisted content constitutes significant original work lies with the authors. In cases of doubt, especially when such tools are used extensively, this should be mentioned in the **general notes**.

- To ensure the **reproducibility** of research results, the inputs (prompts) and outputs of the AI must also be stored for at least 10 years, where necessary. Exceptions apply to data and documents that must be deleted after processing due to specific regulations (e.g., data protection, copyright, terms of use). In such cases, the protection of data, copyrights, etc., takes precedence. Any resulting gaps must be documented.

Personal rights and data protection

Sensitive or protected data must never be uploaded to an online AI system unless the AI Working Group has reviewed and confirmed that the tool in use does not store, share, or reuse this data, for example for training future language models outside of WIFO. In addition, all applicable terms of use, including copyright provisions, must be respected in full. Examples of sensitive or protected data include:

- Trade secrets (e.g., unpublished strategic documents, drafts of forecasts, statements, or press releases),

- Personnel data (e.g., application documents, records of personnel communications),

- Personal information from third parties in emails (e.g., signatures, opinions, personal statements), and

- Unpublished research, particularly if it originates from third parties**.**

The last point refers, for example, to the review of project proposals or manuscripts, where AI models could disseminate and further process them without protecting copyrights.[7]

5. Outlook

The emergence and increasing dissemination of generative artificial intelligence models are transforming scientific and administrative activities and workflows. In light of these developments, WIFO has established a set of guidelines for the use of AI. With this publication, we seek to contribute to the continued evolution of AI governance within the scientific community and

---

[5] For example: "This statement is based on a query to/interaction with ChatGPT on April 6, 2025, https://chat.openai.com."

[6] In WIFO publications, such exceptions may apply, for example, to AI-assisted language corrections and translation services, coding of program scripts, or the use of software for transcribing interviews, provided this does not conflict with the terms of use of the AI tools.

[7] For example, the Deutsche Forschungsgemeinschaft (DFG, 2023) strictly prohibits the use of generative models in preparing reviews due to the confidential nature of review procedures and associated documents.

to foster a collaborative approach to shaping this process. Feedback and the sharing of experiences and ideas from users and other research organisations are highly welcome.

References

ALLEA (2023). Europäischer Verhaltenskodex für wissenschaftliche Integrität – überarbeitete Fassung 2023. Berlin: https://allea.org/wp-content/uploads/2024/09/CoC2023-German.pdf

APA (2023), Leitlinie zum Umgang mit künstlicher Intelligenz. Austria Presse Agentur, Wien.

Birnbauer D. (Hrsg.) (2025). Künstliche Intelligenz in der Praxis, WEKA, Wien, https://www.weka.at/verlag/Datenschutz-IT/Datenschutz/Kuenstliche-Intelligenz-in-der-Praxis-1057030

DFG (2023). Stellungnahme des Präsidiums zum Einfluss generativer Modelle für die Text- und Bilderstellung auf die Wissenschaften und das Förderhandeln der DFG, Deutsche Forschungsgemeinschaft: https://www.dfg.de/resource/blob/289674/ff57cf46c5ca109cb18533b21fba49bd/230921-stellung-nahme-praesidium-ki-ai-data.pdf

EU (2023). Living Guidelines on the Responsible Use of Generative AI in Research, ERA Forum Stakeholder's document, European Commission: https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/guidelines-responsible-use-generative-ai-research-developed-european-re-search-area-forum-2024-03-20_en

EU (2024). Gesetz über Künstliche Intelligenz, Verordnung (EU) 2024/1689 (AI Act), ABl. L 1689, 12.7.2024: https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=OJ:L_202401689

Leibniz Gesellschaft, 2024. Empfehlung zur Sicherung der guten wissenschaftlichen Praxis beim Umgang mit Künstlicher Intelligenz. Beschlossen durch das Präsidium der Leibniz-Gemeinschaft am 29. November 2024: https://www.leibniz-gemeinschaft.de/fileadmin/documents/3085-2906/Leibniz_Empfeh-lung_zur_KI_Beschluss_Praesidium_2024_11_29_.pdf

ORF (2025), KI-Guidelines: glaubwürdig, verlässlich und innovativ, Österreichischer Rundfunk, Wien.

Additional sources

DeeplWrite, ChatGPT and Perplexity AI were used in the creation of this guide for language corrections, translations, specific research tasks (e.g., definitions, citation rules), and for drafting the abstract. The authors have reviewed and edited the content used and take full responsibility for it.