

## Guidelines for the responsible deployment and use of AI tools and technologies and AI research

---

*The University Board of Directors,*

on the basis of art. 2 para. 2, art. 10 para. 3 and art. 39 para. 1 items a and b of the Act governing the University of 5 September 5, 1996 (UniG),

*resolves:*

### 1. Introduction

The present guidelines support the research community at the University of Bern in the responsible deployment and use of Artificial Intelligence (AI) in research. They apply both to the deployment of externally sourced AI tools and to the development and study of AI systems within research projects.

The guidelines aim to ensure that AI-related research activities are aligned with principles of good scientific practice, scientific integrity, and institutional responsibility. They complement existing regulations on research integrity and do not replace applicable legal or institutional requirements.

The principles are grounded in the ethical values and academic mission of the University of Bern and reflect relevant national and international governance frameworks, including the [OECD principles for trustworthy AI](#), Swiss Federal Council's "[Guidelines 'Artificial Intelligence' for the Confederation](#)" from 2020 and, where applicable, the EU AI Act ([Regulation \(EU\) 2024/1689](#)). These guidelines are reviewed yearly under the auspices of the Vice-Rector Research and Innovation.

To support their practical implementation, the University of Bern provides complementary AI literacy materials, including a step-by-step decision guide for researchers, which offer operational guidance and indicate relevant contact points for specific questions or issues.

Furthermore, the University of Bern is providing [practical recommendations on AI](#) tools that may be used within the institutional framework. The use of additional tools, including discipline- or project-specific applications, is permissible provided that applicable legal, information security and data protection requirements are respected.

## 2. Objectives

The aim of the present guidelines is to

- 1) promote the responsible deployment and use of AI in research, including both externally sourced and internally developed systems. This encompasses ethical considerations, transparency, accountability, safety and appropriate human oversight throughout their lifecycle – from conception and design to testing, implementation, and ongoing monitoring. Developers and researchers are expected to adhere to the guiding principles presented in chapter 3.
- 2) promote trustworthy AI research – adhering to the regulations concerning scientific integrity, encouraging reproducibility and openness, while fostering interdisciplinary collaboration and critical reflection on the societal impacts of AI.

## Guiding principles and implementation guidelines

### 3. Guiding principles

The following principles on responsible deployment of AI and trustworthy AI research at the University of Bern provide orientation for researchers and serve as the basis for the implementation of measures set out in Chapter 4.

- AI for the good of humanity

AI must promote the well-being of humans and living beings. It must be compatible with the dignity of life and human autonomy. This includes preserving human skills to support long-term human agency, learning, and resilience.

- Sustainability

AI systems and research should minimize environmental impact and support long-term ecological and social sustainability.

- Inclusivity and diversity

AI systems and research must actively include diverse perspectives and ensure equitable access and outcomes as much as possible. Harm prevention and security

AI systems must be designed, selected, deployed, and used in a way that minimizes foreseeable risks. Appropriate technical and organizational safeguards shall be proportional to the risk profile of the application.

- Methodological suitability and proportionality

The deployment and use of AI must be methodologically appropriate to the specific research objectives and potential risks.

- Transparency and critical review

The deployment and use of AI in research must be transparent and openly communicated. AI-supported processes and outputs must be critically reviewed, contextualized, and communicated together with known limitations and uncertainties.

- Digital sovereignty and technological autonomy

AI development, deployment and use should respect local governance structures, data sovereignty, and the principle of technological autonomy. The deployment, use and integration of AI systems must remain an informed, voluntary, and context-sensitive decision, preserving institutional and individual autonomy over technological choices and digital infrastructures.

- Responsibility and accountability

Human researchers remain accountable for all research processes and outputs involving AI tools, including their integrity, validity, and compliance with applicable legal and ethical standards.

- Data protection, security and privacy

AI systems must uphold the highest standards of data protection, ensuring lawful, secure and privacy-respecting processing of personal and sensitive data, in accordance with applicable institutional and national regulations

- Documentation

AI-related processes, research data, and results must be thoroughly documented to ensure transparency, comprehensibility, and replicability.

#### **4. Implementation of guiding principles**

The following implementation guidelines explain further how to implement the guiding principles presented above (Section 3). By following the implementation guidelines, you follow all the guiding principles as well. Compliance efforts shall be proportionate to the potential impact and risk associated with the respective AI application.

- Transparency

Always disclose the deployment and use of AI in research outputs and documentation, specifying its role within the research process.

Clearly communicate known limitations, uncertainties, and potential methodological impacts of AI deployment and use on your research findings, results and other output to all relevant stakeholders (e.g. peer-reviewers, collaborators, data re-users).

Where applicable, document the type of system used, its version, and relevant configuration parameters.

- Methodological validation

Assess whether AI tools are suitable for the specific research objective.

Critically review AI-generated outputs for factual accuracy and disciplinary validity.

Apply additional validation procedures where AI significantly influences research design, analysis, or conclusions.

Adjust the depth of validation to the potential impact and risk level.

- Risk assessment & bias management

Identify and assess foreseeable technical, methodological, legal, and ethical risks before deployment. Where significant risks cannot be adequately mitigated, avoid the use of AI systems.

Detect and mitigate biases where reasonably possible.

Always reflect on and transparently communicate remaining limitations and residual risks as openly as possible.

- Ethics and accountability

Ensure that AI use complies with existing institutional ethical standards and applicable legal requirements. As a general rule, we recommend incorporating project-specific internal guidelines, as well as existing institutional ethical standards and applicable legal requirements for the ethical deployment and use of AI, into the project agreement.

Conduct ethics reviews where research involves potential impacts on sensitive data, vulnerable groups, or significant societal impact.

Human researchers remain fully accountable for the integrity, originality, and validity of research involving AI.

Outputs generated with generative AI must be reviewed for factual accuracy, contextual appropriateness, and disciplinary validity by a qualified human expert prior to publication or dissemination. Researchers must be transparent about the deployment and use of AI and the prompts that were given.

AI systems cannot qualify as authors or co-authors.

- Technological autonomy and digital sovereignty

Assess potential dependencies on proprietary infrastructures and vendor lock-in effects that may affect long-term accessibility, reproducibility, and institutional control over research processes.

Prefer solutions that allow adequate control over data, documentation, and reproducibility where feasible.

Periodically review whether continued deployment and use of a system remain aligned with institutional governance standards and research values.

- Data protection

Implement appropriate technical and organizational security measures and ensure compliance with institutional, cantonal, and national data protection requirements.

Always check intellectual property rules and contractual obligations (for instance with publishers or grant providers) before uploading content into external AI systems.

Anonymize or pseudonymize data wherever required.

Ensure lawful processing, informed consent where required, and develop clear data use agreements when collaborating with external partners.

- Human oversight

Ensure human oversight in any critical decisions and processes involving AI.

Critically evaluate AI outputs in light of disciplinary expertise.

Establish mechanisms for documentation, correction and traceability appropriate to the system's impact and risk profile.

- Inclusive use

Conduct proportionate impact assessments that consider potential effects on individuals and groups with diverse social and cultural backgrounds.

Engage relevant stakeholders where appropriate to identify risks of exclusion, discrimination, or unequal access.

Seek fair and transparent balancing solutions where competing interests arise.

Periodically review whether AI-supported processes generate unintended exclusionary effects.

## **5. Breaches of scientific integrity**

The implementation instructions as well as the guiding principles are part of good scientific practice. Therefore, violations of the implementation instructions as well as the guiding principles may constitute breaches of the regulations concerning scientific integrity and will be handled in accordance with those regulations.

## **6. Entry into force**

These guidelines are effective following the decision of the University Board of Directors

Bern, April 28, 2026

On behalf of the University Board of Directors



Prof. Dr. Virginia Richter

Rector

Appendix

[Responsible use of GenAI: A Step-by-Step Guide](#)