The impact of generative models for text and image creation on good research practice and on research funding

Dr Sonja Ochsenfeld-Repp | German Research Foundation

Oral Presentations 5 | June 3, 2024 | WCRI | Athens, Greece

Generative models for text and image creation
Setting the scene

► Text and image creation / Large Language Models (Generative AI)

► Wide-ranging impact of generative models for text and image creation

- changing work processes in science and the humanities, knowledge production, creativity

► Epistemic dimensions for research as well as for research integrity.

► Recently (2023), Nature conducted a survey of 1.600 researchers on the benefits of AI, https://www.nature.com/articles/d41586-023-02980-0 *(Van Noorden/Perkel).*

► Cornerstones of research integrity: research process and findings have to be transparent and comprehensible to third parties → holds true for applications to funding organisations.

► Professional ethics of researchers: full compliance with good research practice. The use of generative models cannot relieve researchers of this content-related and formal responsibility.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

2   Athens, 3 June 2024

DFG

► September 2023, initial Statement by the DFG´s Executive Committee on "The influence of generative models for text and image production on the sciences and humanities and on the DFG's funding activities", https://www.dfg.de/resource/blob/289676/89c03e7a7a8a024093602995974832f9/230921-statement-executive-committee-ki-ai-data.pdf

► April 2024, key points formulated in the statement were incorporated into the proposal guidelines and specified in the Research Integrity Portal, https://wissenschaftliche-integritaet.de/en/use-generative-models/

- Applicants must disclose the use of generative models in the application (if so, which ones, for what purpose and to what extent). The content can be described in a few explanatory sentences; not necessary to identify the text passages.

- AI tools used that do not affect the scientific content of the application (e.g. grammar, style, spell-checking, translation programs) do not have to be specified.

- In decision-making processes, the use of generative models by applicants is assessed to be neither positive nor negative.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

3   Athens, 3 June 2024

DFG

► Research Funders Policy Group (e.g. Wellcome, UKRI), U.S. National Science Foundation, National Institutes of Health, etc., permit the use of generative models in applications, but this must be clearly acknowledged.

► The EU "Living Guidelines on Responsible Use of AI in research", published in March 2024, are also moving in this direction, https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf.

► Only natural persons responsible can act as applicants. They must ensure that the use of generative models does not infringe anyone else's intellectual property and does not result in scientific misconduct.

- With generative models, which do not adopt text passages verbatim but work on the basis of probabilities, copyright infringement (exploitation) is ruled out in most of the cases, https://hss-opus.ub.ruhr-uni-bochum.de/opus4/frontdoor/deliver/index/docId/9734/file/2023_03_06_Didaktik_Recht_KI_Hochschulbildung.pdf

- Unidentified use of generative models to create text/images in proposals: Concealment that AI has been used is not misrepresentation (falsification and fabrication of data); plagiarism is a possibility.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

4   Athens, 3 June 2024

DFG

▶ Due to legal uncertainties and challenges, the use of generative models in the preparation of reviews for applications is currently largely classified internationally as inadmissible or associated with considerable risks.

- National Institutes of Health, National Science Foundation, Australian Research Council, Swiss National Science Foundation, Research Funders Policy Group, etc., prohibit the use of generative models.

- Amongst those Science Europe member organisations that have published guidelines/recommendations on the use of generative models at all, the use of such tools that infringe applicants' confidentiality are illicit.

- The respective DFG´s rules are consistent with this.

▶ The reason for the current exclusion is to safeguard the integrity of the funding procedures.

- Confidentiality of the assessment process (regarding written and oral content) is a particularly valuable asset.

- Unpublished research ideas and information relevant to copyright/patent law have to remain in the realm of the funder.

- (Personal) data is transmitted to the funding organisation solely for the purpose of checking eligibility and assessing the proposed research.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

5   Athens, 3 June 2024

DFG

► The "EU Living Guidelines" allow the use of generative models under certain, yet underdetermined circumstances. They transfer responsibility for use/protection of data and confidentiality to the users of the AI.

- "[Researchers] (…) protect unpublished or sensitive work (such as their own or others' unpublished work) by taking care not to upload it into an online AI system unless there are assurances that the data will not be re-used (…). Researchers take care not to provide third parties' personal data to online generative AI systems unless the data subject (individual) has given them their consent and researchers have a clear goal for which the personal data are to be used so compliance with EU data protection rules. (…)"

► The input of application content into generative models is currently comparable to the unauthorised disclosure of an application to a third party by an expert. This may constitute:

- Breach of confidentiality (violation of good research practice, potentially research misconduct)
- Copyright infringement regarding the texts of the proposal (potentially used as training data/ reproduction)
- Breach of data protection law in the case of personal data.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

6   Athens, 3 June 2024

DFG

► Use of generative models must be considered in light of the EU AI ACT adopted by the Council of Europe on May 21, 2024, https://artificialintelligenceact.eu/.

- Which risk class defined by the law, with which consequences for handling, does research assessment carried out with the help of generative models fall into?

- Scope of application of law will generally be opened up when generative models are used to create expert opinions, as so-called general-purpose AI will be used (not pure research software, which is excluded from the scope of the law).

- Standard examples of high-risk use include the following "education and vocational training" and "access to and enjoyment of essential private services and essential public services", Annex III No. 3 and No. 5.

► Current status of handling may change as result of technical possibilities, e.g. on-site/ on-device systems.

► Fundamental research ethics question: Should generative models be used in the peer review process at all?

- Numerous questions remain unanswered, yet, e.g. the handling of corresponding expert opinions in funding decisions, the fundamental ability of generative models to evaluate research in a science-led and unbiased manner.

The impact of generative models for text and image creation on good research practice and on research funding | Ochsenfeld-Repp

7   Athens, 3 June 2024

DFG

Dr Sonja Ochsenfeld-Repp

Head of Division "Research Culture"

sonja.ochsenfeld-repp@dfg.de

For more information

► on the DFG: **www.dfg.de/en**
► on DFG-funded projects: **www.gepris.dfg.de/en**
► on the online Portal "Research Integrity": **https://wissenschaftliche-integritaet.de/englisch/**
► on over 23,000 German research institutes: **https://gerit.org/en/**