

Workshop

Text-generating AI and good research practice

27 June 2023, 1- 6pm; Berlin-Brandenburgische Akademie der Wissenschaften

I.

The German Research Ombudsman organised its first one-day workshop on 27 June 2023 on the topic of text-generating artificial intelligence (AI) and good research practice (GRP), as part of the “Discussion Hubs to Foster Research Integrity” project. The mutual exchange centred on questions of how text-generating AI can be applied in research in conformity with GRP, and whether additional recommendations or guidelines on the subject are necessary.

II.

The event was prompted by recent developments in generative AI. ChatGPT (based on GPT 3.5) was released at the end of November 2022. Several research institutions reacted quickly by issuing warnings, and banned its use. Some research journals and scientific publishers released statements on how to approach text-generating AI. With regard to some points an initial consensus was quickly reached: Neither text-generating AI nor ChatGPT is able to meet the requirements for authorship, and its use must be made transparent.

However a large number of unanswered questions remain, especially regarding how text-generating AI is to be used for research publications, as well as other issues that have only been superficially resolved. These include currently ambiguous details of how to use text-generating AI in conformity with GRP.

It is also largely unclear how – once rules on AI and GSP have been established – any breaches of these rules are to be dealt with. At present, there is virtually no reliable way to verify and prove that text-generating AI has been used. Some commentators see the situation as a technological arms race: all that needs to be done is to develop powerful “recognition software” in order to detect the undesirable use of text-generating AI. Such detection tools, which are already being intensively advertised by commercial providers, however appear limited in terms of their reliability.

III.

Seven invited experts attended the workshop:

- **Dr Katharina Beier**, Head of the Ombuds Office of the University of Göttingen
- **Prof. Dr Iryna Gurevych**, Director of the Ubiquitous Knowledge Processing Lab at the Technical University of Darmstadt’s Computer Science Department, and Head of the “Artificial Intelligence for living texts” project

- **Dr Guido Juckeland**, Head of Computational Science and Ombudsperson at the Dresden-Rossendorf Helmholtz Centre
- **Dr Kirsten Hüttemann**, Director of the Research Culture Group at the German Research Foundation (DFG), responsible amongst other things for the implementation of the DFG's Code of Conduct and for the third level of the Code of Conduct
- **Nadine Lordick**, Member of staff at the University of Bochum's Centre for Teaching and Learning and on the KI:edu.nrw project
- **Prof. Dr Debora Weber-Wulff**, Emeritus Professor of Media and Computing at Berlin University of Applied Sciences
- **Prof. Dr Doris Weißels, Professor for Business Information Systems** at Kiel University of Applied Sciences, with a **research focus** on Natural Language Processing

The following members of the Committee and of the Office of the *German Research Ombudsman* attended:

- **Renate Scheibe**, University of Osnabrück, member of the Committee
- **Daniela N. Männel**, University of Regensburg, member of the Committee
- **Roger Gläser**, University of Leipzig, member of the Committee
- **Katrin Frisch**, project coordinator in the discussion hub "dealing with data"
- **Felix Hagenström**, project coordinator in the discussion hub "dealing with plagiarism"
- **Nele Reeg**, project coordinator in the discussion hub "dealing with authorship conflicts"
- **Hjördis Czesnick**, Head of the Office of the *German Research Ombudsman*
- **Sophia May**, Research Integrity Advisor at the Office of the *German Research Ombudsman*

IV.

The question as to how to deal with text-generating AI in research relates to two aspects: Firstly, the fundamental values as well as general understanding of science and the professional ethos. What does it mean to work *lege artis* in light of the changes brought about by text-generating AI? How should we interpret responsibility and reliability? What is a scientific text intended to achieve? Which functions does authorship fulfil in a scientific context? The second aspect relates specifically to GRP and the proper conduct of research. These two levels are naturally interlinked; yet the workshop primarily focussed on the level of concrete practice, with emphasis being placed on scientific research (rather than teaching). The discussion was orientated towards key questions clustered around four topics:

- 1) "Authorship and evaluation practices"
- 2) "Transparency"
- 3) "Deviations from GRP", and
- 4) "Vision 2030 – the 'new normal' with text-generating AI in science".

During the workshop, it became increasingly clear that a consensus already exists on many of the questions given. As a next step, productive discussions on how to deal with text-generating AI in science need to be facilitated in the disciplines themselves, as this is the only place where diverging disciplinary conventions, authorship criteria, and different notions of text could be suitably addressed. Recommendations across disciplines beyond existing guidelines included in the DFG's Code of Conduct were regarded as being less helpful. Values such as honesty and integrity that are set out in the Code, as well as the requirement for sources to be suitably identified, would become no less valid were new technologies to be included. However, specific practices need to be adjusted in line with the new circumstances and existing guidelines interpreted for the new context via additional statements.

A topic intensely debated in the workshop was the issue of transparency, specifically how the use of text-generating AI could best be made transparent in publications. There were divergent opinions as to whether it was sufficient in terms of transparency to specify the programmes that were being used (ideally also stating the precise version), and the purpose to which they were being put, or whether it was necessary – as some journals are also already demanding – to also chronicle and submit the underlying prompts. With regard to the latter, it needs to be taken into account that working with text-generating AI usually happens in stages, and in a sense collaboratively. Submitting all prompts used for a publication does not do justice to this *modus operandi*, or might even be a requirement impossible to fulfil. In some cases the authors' own intellectual contribution is difficult to determine, especially where tools are used which provide support both in drafting texts, evaluating research literature as well as in designing the research question – as, for example, the programme Elicit does. This illustrates that there is a need to take into consideration differences among disciplines when formulating AI-specific rules for referencing and citation.

The workshop made it clear that regarding transparency, conventions need to emerge in synch with the rapid, ongoing dynamic developments in artificial intelligence, where future changes could be anticipated only to a limited degree. For now, scholars and research institutions need to tolerate current uncertainties.

In addition to taking a closer look at discipline-specific differences, the workshop addressed the characteristics of different text genres. Besides the use of AI generated text in “conventional” research literature, attention also needs to be paid to its use with abstracts, grant applications and reviews. The question of what motivates researchers to use AI plays a role here. Drafting functional texts such as abstracts using text-generating AI does not necessarily give rise to serious suspicions of deception. The practice of drawing up abstracts, however, also serves a purpose which goes further than simply writing a summary. Composing different types of text helps researchers to gain expertise in scientific writing. The value accorded to this ability, in turn, depends greatly on the academic context. In the workshop, it was rather the matter of (peer) reviews generated by AI that was regarded as more problematic. What needs to be averted is the creation of an absurd system in which text-generating AI was responsible for both producing *and* evaluating texts. Also, similarly to using software to detect plagiarism, detecting AI-generated texts via software is unreliable, as Deborah Weber-Wulff has made clear in [her recent study](#).

The question of deception through the use of AI was further debated. There is a need to clarify what we mean by “deception” with regard to text-generating AI. Is text-generating AI generally based on the principle of deception? Who is deceiving whom, about what, by using text-generating AI (and to what end)? Does a lack of transparency always constitute deception? These questions need to be answered, once more with regard to different text genres, given that deception is context sensitive.

Discussing deception, the workshop opened up a critical perspective vis-à-vis problems that are inherent in present-day academia. Using text-generating AI accelerates the already fast moving pace within science, which operates according to the maxim of “publish or perish”. The question of “Why deceive?”, raised in the workshop, thus led to a discussion about the system of incentives which rewarded the quantity of scientific output rather than its quality, all this alongside general time and performance pressures. Text-generating AI could act here as a kind of negative super booster, without being the trigger or cause of these problems itself. The attention currently devoted to ChatGPT and other programmes should equally be directed towards the problems already present within the system. At the same time, the participants also stressed that Germany, and the university governing bodies in particular, had a great deal of catching up to do when it comes to AI. Even though there was considerable interest in the topic, the university governing bodies often lacked information, insight as well as guidance on how to deal with text-generating AI in teaching and research. This includes keeping an eye on not only the use of text-generating AI, but also its development. The most powerful text-generating AI systems at the moment are proprietary. This means that the development is not only subject to commercial exploitation, but it can also lead to an imbalance in terms of who can afford to use powerful text-generating AI systems in the future. The workshop showed very clearly that reflecting on text-generating AI also means thinking about enduring problems with which the scientific community is continuously confronted.

A paper on the topic incorporating the results of the workshop is to be included in the special issue on good research practice of the *Zeitschrift für Bibliothekswesen und Bibliographie (ZfBB)*, due to be published at the end of 2023.