

Workshop

Textgenerierende KI und gute wissenschaftliche Praxis

27. Juni 2023, 13:00-18:00 Uhr; Berlin-Brandenburgische Akademie der Wissenschaften

I.

Am 27. Juni 2023 veranstaltete der Ombudsman für die Wissenschaft im Rahmen der „Dialogforen zur Stärkung der Kultur wissenschaftlicher Integrität“ erstmalig einen eintägigen Workshop zum Thema textgenerierende künstliche Intelligenz (TKI) und gute wissenschaftliche Praxis (GWP). Der gemeinsame Austausch widmete sich den Fragen, wie der GWP-konforme Umgang mit TKI in der Wissenschaft aussehen kann und ob zusätzliche GWP-Empfehlungen notwendig sind.

II.

Anlass und Hintergrund der Veranstaltung waren die jüngsten Entwicklungen auf dem Feld der TKI. Ende November 2022 wurde ChatGPT (basierend auf GPT 3.5) veröffentlicht. Ziemlich schnell reagierten manche Wissenschaftsinstitutionen mit Warnungen und Nutzungsverböten. Einige Fachzeitschriften und Wissenschaftsverlage veröffentlichten Stellungnahmen zum Umgang mit TKI. Hier war man sich zumindest in einigen Punkten bald recht einig: TKI wie ChatGPT kann die Voraussetzungen für Autorschaft nicht erfüllen und ihre Nutzung muss transparent gemacht werden.

Viele Fragen zum Einsatz von TKI bei der Erstellung von wissenschaftlichen Veröffentlichungen wurden jedoch bisher nicht oder nur oberflächlich beantwortet. Wie die GWP-konforme Nutzung von TKI im Rahmen von Wissenschaft und Forschung im Detail aussehen kann, bleibt vorerst unklar.

Ebenfalls weitgehend unklar ist, wie – nach der ersten Festlegung von GWP-konformen Nutzungsregeln – mit möglichen Regelverstößen umgegangen werden soll. Der Einsatz von TKI lässt sich kaum zuverlässig überprüfen und nachweisen. Manchmal ist nun auch von einem technologischen „Wettrüsten“ die Rede: Man müsse nur eine potente „Erkennungssoftware“ entwickeln, um den ungewünschten Einsatz von TKI aufzudecken. Die Zuverlässigkeit solcher Detektionstools, für die bereits fleißig durch kommerzielle Anbieter geworben wird, scheint jedoch begrenzt.

III.

An dem Workshop nahmen sieben geladene Expert:innen teil:

- **Dr. Katharina Beier**, Leiterin der Ombudsstelle der Universität Göttingen
- **Prof. Dr. Iryna Gurevych**, Direktorin des Ubiquitous Knowledge Processing Lab am Fachbereich Informatik an der TU Darmstadt und Leiterin des Projekts „Artificial Intelligence for living texts“
- **Dr. Guido Juckeland**, Leiter Computational Science und Ombudsperson am Helmholtz-Zentrum Dresden-Rossendorf

- **Dr. Kirsten Hüttemann**, Direktorin der Gruppe Forschungskultur der DFG und unter anderem zuständig für die Umsetzung des DFG-Kodex und die Dritte Ebene des Kodex.
- **Nadine Lordick**, Mitarbeiterin am Zentrum für Wissenschaftsdidaktik der Universität Bochum und im Projekt KI:edu.nrw
- **Prof. Dr. Debora Weber-Wulff**, emeritierte Professorin für Informatik an der Hochschule für Technik und Wirtschaft Berlin
- **Prof. Dr. Doris Weßels**, Professorin für Wirtschaftsinformatik an der Fachhochschule Kiel mit dem Forschungsschwerpunkt Natural Language Processing

Seitens des Gremiums und der Geschäftsstelle des *Ombudsman für die Wissenschaft* nahmen teil:

- **Renate Scheibe**, Universität Osnabrück, Mitglied des Gremiums
- **Daniela N. Männel**, Universität Regensburg, Mitglied des Gremiums
- **Roger Gläser**, Universität Leipzig, Mitglied des Gremiums
- **Katrin Frisch**, Referentin im Dialogforen-Projekt zum Thema Daten
- **Felix Hagenström**, Referent im Dialogforen-Projekt zum Thema Plagiate
- **Nele Reeg**, Referentin im Dialogforen-Projekt zum Thema Autorschaften
- **Hjördis Czesnick**, Leiterin der Geschäftsstelle des *Ombudsman für die Wissenschaft*
- **Sophia May**, Mitarbeiterin in der Geschäftsstelle des *Ombudsman für die Wissenschaft*

IV.

Die Frage nach dem Umgang mit TKI in der wissenschaftlichen Forschung betrifft zweierlei. Zum einen die grundsätzlichen Werte wissenschaftlichen Arbeitens, das Berufsethos und das allgemeine Verständnis von Wissenschaft. Was bedeutet *lege artis zu arbeiten* angesichts der Veränderungen durch TKI? Wie sind Verantwortung und Zuverlässigkeit auszulegen? Was soll ein wissenschaftlicher Text leisten? Welche Funktion hat die Verbindung von Autorschaft und Text im wissenschaftlichen Kontext? Zum anderen sind die konkrete GWP bzw. die Regeln guten wissenschaftlichen Arbeitens betroffen. Diese zwei Ebenen sind natürlich miteinander verschränkt; im Workshop wurde sich hauptsächlich auf die Ebene der konkreten Praxis konzentriert. Der Fokus lag dabei auf der wissenschaftlichen Forschung (und nicht der Lehre). Die Diskussion orientierte sich an Leitfragen, die vier Themenbereichen zugeordnet waren:

- 1) „Autorschaft und Bewertungspraktiken“
- 2) „Transparenz“
- 3) „Abweichungen von der GWP“ sowie
- 4) „Vision 2030 – die „neue Normalität“ mit TKI in der Wissenschaft“.

Im Verlauf des Workshops wurde festgestellt, dass zu vielen Fragen bereits weitgehend Konsens besteht. So wurde deutlich, dass eine produktive Diskussion zum Umgang mit TKI in der Wissenschaft am besten in den Fachcommunities selbst geführt werden sollte, da nur dort divergierende Fachgepflogenheiten, Autorschaftsdiskurse sowie Textverständnisse adäquat adressiert werden könnten. Fachübergreifende Empfehlungen, die über die bestehenden Leitlinien des DFG-Kodex hinausgehen, um größeren Fokus auf das Thema KI zu legen, wurden als weniger hilfreich eingeschätzt.

Denn kodifizierte Werte wie Ehrlichkeit, Redlichkeit sowie die Forderung der genügenden Ausweisung von Quellen verlören durch den Einbezug neuer Technologien nicht an Gültigkeit. Jedoch bestehe Bedarf, gewisse Praktiken an die neuen Gegebenheiten anzupassen sowie bestehende Leitlinien durch zusätzliche Erläuterungen für den neuen Kontext auszulegen.

So wurde zum Beispiel im Workshop intensiv diskutiert, wie die Nutzung von TKI in Veröffentlichungen am besten transparent gemacht werden könne. Dabei gingen die Meinungen auseinander, ob es als Mindestanforderung an die Transparenz ausreiche die verwendeten Programme (idealerweise inklusive der Angabe der genauen Version) und ihren Verwendungszweck anzugeben, oder ob – wie bereits auch von einigen Journalen gefordert – die verwendeten Prompts ebenso anzugeben seien. Zur zweiten Variante wurde zu bedenken gegeben, dass das Arbeiten mit TKI mehrschrittig und kollaborativ sei, so dass letztere Option dieser Arbeitsweise nicht gerecht werde oder sogar unmöglich zu erfüllen sei. Teilweise sei die intellektuelle Eigenleistung nicht so klar abgrenzbar, insbesondere wenn Tools zum Einsatz kommen, die neben der Textgenerierung auch bei der Auswertung von Forschungsliteratur und der Konzeption der Forschungsfrage – wie beispielsweise das Programm Elicit – unterstützen. Es zeigte sich an diesem Beispiel, dass Disziplinunterschiede bei der Aufsetzung von Regeln zur Kenntlichmachung mitgedacht werden müssen.

Ebenso wurde in Bezug auf Transparenz im Workshop deutlich, dass Konventionen sich erst herausbilden müssen, und zwar bei einer anhaltenden dynamischen und rasanten Entwicklung im Bereich der künstlichen Intelligenz, dessen zukünftige Veränderungen nur begrenzt antizipiert werden können. So gehe es für den Forschungsbetrieb jetzt auch darum, die gegenwärtigen Unsicherheiten aushalten zu können.

Neben dem Blick auf disziplinspezifische Unterschiede kamen auch immer wieder die Charakteristika verschiedener Textgattungen zu Sprache. So ist neben dem Einsatz von TKI bei „klassischer“ Forschungsliteratur auch der Einsatz bei Abstracts, Förderanträgen und Reviews zu beachten. Dabei spielt die Frage eine Rolle, aus welcher Motivation heraus Forschende TKI nutzen. Funktionstexte wie Abstracts könnten beispielsweise ohne Probleme mithilfe von TKI erstellt werden, ohne dass dies den schwerwiegenden Verdacht von Täuschung aufwerfen müsste. Jedoch dient auch die Praktik des Abstract-Schreibens einem Zweck, der über die reine Erstellung einer Zusammenfassung hinausgeht. So schult das Verfassen verschiedener Textgattungen das wissenschaftliche Schreiben. Welcher Wert dieser Fähigkeit beigemessen wird, hängt dabei wieder stark vom Fachkontext ab. Im Workshop wurden denn auch von TKI automatisch verfasste Reviews/Gutachten als problematischer beurteilt. Insbesondere sei vor einem absurden System zu warnen, in dem TKI gleichzeitig maßgeblich für die Produktion *und* Bewertung von Texten zuständig ist. Die Detektion von KI-generierten Texten mittels Software ist – analog zur Plagiatsermittlung durch Software – unzuverlässig, wie Deborah Weber-Wulff mit einer kürzlich durchgeführten Studie [<https://doi.org/10.48550/arXiv.2306.15666>] verdeutlichte.

Es wurde zudem die Frage nach der Täuschung durch den Einsatz von TKI aufgeworfen. Dabei muss der Begriff der Täuschung in Bezug auf TKI differenziert betrachtet werden. Beruht TKI generell auf dem Prinzip der Täuschung? Wer täuscht wen über was durch die Nutzung von TKI (und zu welchem Zweck)? Ist fehlende Offenlegung stets eine Täuschung? Diese Fragen gilt es auch wieder in Bezug auf die unterschiedlichen Textgattungen zu beantworten, denn Täuschung ist kontextsensitiv.

Insbesondere auch über die Thematik der Täuschung eröffnete der Workshop eine kritische Perspektive auf bestehende Problematiken im gegenwärtigen Wissenschaftssystem. Der Einsatz von TKI beschleunige den ohnehin schnelllebigen Wissenschaftsbetrieb, der unter der Maxime „publish or perish“ operiert. So führte die Frage „Warum wird getäuscht?“ auch zum Anreizsystem, das die Quantität mehr als die Qualität des wissenschaftlichen Outputs belohne. Hinzu kämen allgemeiner Zeit- und Leistungsdruck. TKI könne hier als eine Art negativer Super Booster wirken, ohne selbst Auslöser bzw. Grund dieser Probleme zu sein. Die Aufmerksamkeit, die zurzeit ChatGPT und anderen

Programmen gewährt wird, sollte ebenso auf die schon bestehenden systemimmanenten Probleme gerichtet werden. Gleichzeitig betonten die Teilnehmer:innen auch, dass Deutschland und insbesondere auch die Hochschulleitungen einen großen Nachholbedarf in Sachen KI hätten. Obwohl das Interesse am Thema groß sei, fehlten von Seiten der Hochschulleitungen oft Informationen, Aufklärungen und Regelungen zum Umgang mit TKI in Lehre und Forschung. Dazu gehöre ebenfalls nicht nur die Nutzung, sondern auch die Entwicklung von TKI im Blick zu behalten. Momentan sind die leistungsfähigsten TKI proprietär. Dies bedeutet, dass die Entwicklung nicht nur kommerziellem Nutzen untersteht, sondern es auch zu einem Ungleichgewicht führen kann, wer sich in Zukunft die Nutzung potenter TKIs leisten kann. Nachdenken über TKI, das zeigte der Workshop in aller Deutlichkeit, ist auch ein Nachdenken über anhaltende Probleme in der Wissenschaft.

Eine Publikation zum Thema, die sich aus den Ergebnissen des Workshops speist, ist für das Themenheft „Gute wissenschaftliche Praxis“ der *Zeitschrift für Bibliothekswesen und Bibliographie (ZfBB)* geplant. Dieses erscheint voraussichtlich Ende 2023.