

L 02

KI-Korrekturhilfe von Fobizz bei Evaluation durchgefallen – wie bewertet der Senat die Ergebnisse?

Anfrage der Abgeordneten Dr. Franziska Tell, Dr. Henrike Müller und Fraktion BÜNDNIS 90/DIE GRÜNEN

Wir fragen den Senat:

1. Wie bewertet der Senat die Studienergebnisse der Universität Osnabrück zu der KI-Korrekturhilfe des Unternehmens Fobizz, welche zum Teil auch von Bremer Lehrkräften genutzt wird?
2. Welche Möglichkeiten haben Schüler:innen zu erkennen, ob die Bewertung ihrer Arbeit durch ein von der Lehrkraft genutztes KI-Tool vorgenommen wurde und welche Beschwerdemöglichkeiten gibt es in diesem Fall?
3. Was plant der Senat, um sicherzustellen, dass in Schulen genutzte KI-Tools didaktisch geeignet sind und systematisch evaluiert werden?

Zu Frage 1:

Der Senat hat die besagte Studie „Chatbots im Schulunterricht: Wir testen das Fobizz-Tool zur automatischen Bewertung von Hausaufgaben“ von Rainer Mühlhoff und Marte Henningsen zur Kenntnis genommen.

Die Studie identifiziert signifikante Mängel bei der untersuchten „KI-Korrekturhilfe“ von Fobizz. Zu den zentralen Ergebnissen zählen unter anderem eine hohe Zufälligkeit und Inkonsistenz der Bewertungen und des Feedbacks, eine unzuverlässige Erkennung von Fehlern und Nonsens-Abgaben sowie die Feststellung, dass die Einarbeitung der Verbesserungsvorschläge nicht zu besseren Ergebnissen führt. Die Autoren führen die Defizite auf grundlegende Limitationen der KI-Technologie zurück und schlussfolgern, dass das Tool für den Einsatz im Schulalltag ungeeignet ist und seine Vermarktung irreführend sei.

Dem Senat liegen keine Erkenntnisse darüber vor, ob und falls ja, in welchem Umfang speziell die „KI-Korrekturhilfe“ von Bremer Lehrkräften eingesetzt wird. Es ist bekannt, dass die Fobizz KI-Tools, vor allem die enthaltenen geschützten KI-Chaträume für Schüler:innen, an einigen Schulen genutzt werden.

Lehrkräfte können sich z. B. bei der Formulierung von Feedback oder der Erstellung von Erwartungshorizonten technisch unterstützen lassen. Grundsätzlich gilt aber im Land Bremen, dass die Beurteilung und Bewertung von Schülerinnenleistungen in der Verantwortung der Lehrkräfte liegt und nicht an Dritte – auch nicht an KI-Systeme – delegiert werden darf.

Zu Frage 2:

Wie bereits dargelegt, ist die Delegation der Leistungsbewertung an KI-Systeme in Bremen unzulässig. Schüler:innen müssen darauf vertrauen können, dass ihre Lehrkräfte die Bewertungen selbstständig und nach pädagogischen Grundsätzen vornehmen.

Sollte dennoch der begründete Verdacht bestehen, dass eine Bewertung unzulässigerweise durch ein KI-Tool erfolgt ist, stehen den Schüler:innen sowie deren Erziehungsberechtigten die üblichen Beschwerdewege offen. Technische Lösungen zur Erkennung von KI generierten Texten liefern nur Wahrscheinlichkeitswerte und sind nicht hinreichend zuverlässig.

Zu Frage 3:

Der Senat begleitet die Entwicklung und das Angebot von KI-Tools für den Bildungsbereich sehr aufmerksam. KI-gestützte Tools, die von der Senatorin für Kinder und Bildung zentral bereitgestellt werden, werden vorab auf ihre Eignung und ihre datenschutzrechtliche Unbedenklichkeit überprüft. Die Verantwortung für den Einsatz weiterer KI-gestützter Software liegt, wie bei allen Lehr- und Lernmitteln, in der Verantwortung der jeweiligen Lehrkraft, bzw. der Schule. Eine fachliche Einschätzung oder die systematische Evaluation aller KI-Tools am Markt ist in Anbetracht der Vielzahl verfügbarerer Produkte nicht leistbar. Um die Lehrkräfte bestmöglich zu unterstützen, befindet sich aber gerade eine Handreichung zum KI-Einsatz in Schule in der finalen Abstimmung. Sie beinhaltet formale, rechtliche und praktische Hinweise.