

Aktuelles zu digitalen Prüfungen – welche Zukunft hat KI?

Bericht über die Tagung des Vereins zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. am 18.4.2024

Bei all der Euphorie um die viel beschworenen Fähigkeiten künstlicher Intelligenz (KI) war es wohl nur eine Frage der Zeit, bis die Technologie Einzug hält in die Hörsäle dieses Landes. In einer nicht repräsentativen Umfrage des Digitalverbands Bitkom e.V. gaben im Januar dieses Jahres 65 % der Studierenden an, das generative KI-System ChatGPT zu nutzen. Mehr als die Hälfte der Befragten unterstützte außerdem die These, dass sich Studierende durch den Einsatz des KI-Systems einen ungerechten Vorteil verschaffen können. Dennoch adressiert nur ein kleiner Teil der Prüfungsordnungen deutscher Hochschulen explizit den Einsatz künstlicher Intelligenz.¹ Braucht es also eine umfassende Reform der universitären Regelwerke? Und worauf haben Hochschulen zu achten, wenn sie die Technologie zulassen oder verbieten wollen? Um diesen Fragen nachzugehen, hat der Verein zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. am 18.4.2024 eine Tagung unter der Überschrift „Aktuelles zu Digitalen Prüfungen – Welche Zukunft hat KI?“ veranstaltet. Mehr als 180 Teilnehmerinnen und Teilnehmer folgten der Einladung und verdeutlichten so das rege Interesse an der Thematik. Unter der Moderation des Ehrenvorsitzenden des Vereins, *Prof. Ulf Pallme König*, und des Vorstandsmitglieds *Prof. Dr. Volker Epping* erläuterten zunächst *Edgar Fischer*, Vorsitzender Richter am VG Berlin, und *Dr. Peter Dieterich, LL.M.*, Richter am VG Berlin, die prüfungsrechtlichen Herausforderungen des KI-Einsatzes (I.).² Im Anschluss widmete sich *Prof. Dr. Rolf Schwartmann*, Leiter der Kölner Forschungsstelle für Medienrecht an der TH Köln, den KI-rechtlichen Voraussetzungen einer Implementierung der Technologie an Hochschulen (II.), bevor die Referenten in einer abschließenden Diskussionsrunde auf die offenen Fragen der Teilnehmerinnen und Teilnehmer eingingen (III.).

I. KI-Einsatz in Prüfungen

Den Auftakt zur Veranstaltung machten *Fischer* und *Dieterich* mit einer Betrachtung der rechtlichen Probleme des Einsatzes künstlicher Intelligenz in Prüfungssituationen. Dabei knüpften sie an die vier vorangegangenen Veranstaltungen des Vereins zu Digitalen Prüfungen an, die als Reaktion auf die Unsicherheiten in der Corona-Pandemie bereits im Oktober 2020 eingeführt wurden und sich seither stetig wachsender Beliebtheit erfreuen.³

1. KI-Einsatz seitens des Prüflings

Die in den vorangegangenen Veranstaltungen vermittelte Täuschungsdogmatik im Prüfungsrecht lasse sich auf den Einsatz künstlicher Intelligenz seitens der Prüflinge übertragen. Eine sanktionierbare Täuschungshandlung liege demnach vor, wenn ein Prüfling eine KI-Anwendung ohne Angabe der Nutzung einsetzt. In der Regel sei der verdeckte Einsatz künstlicher Intelligenz als unerlaubte Hilfe im Sinne der Prüfungsordnungen zu qualifizieren. Die Regelwerke müssten daher nicht grundlegend geändert werden. Verboten sei allerdings nur die Ausgabe einer KI-generierten Leistung als eigene, etwa wenn sich der Prüfling eine Hausarbeit von einer KI schreiben lässt. Der KI-Einsatz zu Recherchezwecken dürfte hingegen vorbehaltlich spezieller – kaum effektiv durchsetzbarer – Regelungen in den Prüfungsordnungen oder Eigenständigkeitserklärungen keinen Bedenken begegnen, da die Verwendung von KI in diesem Zusammenhang nicht anders zu bewerten sein dürfte als gebräuchliche Recherchemittel wie das Internet. Zwar seien Antworten von ChatGPT mangels geistiger Eigenleistung einer identifizierbaren Person keine klassische zitierfähige Quelle i.S.d. Urheberrechts – Maßstab des Prüfungsrechts sei jedoch die Eigenständigkeit der Prüfungsleistung, die sich durch Kennzeichnung der Übernahme KI-generierter Bausteine durchaus belegen lasse. Eine solche Offenlegung begegne angesichts der fehlenden

¹ <https://www.bitkom.org/Presse/Presseinformation/So-digital-sind-Deutschlands-Hochschulen> (26.4.2024).

² *Fischer* und *Dieterich* sind Autoren des Standardwerkes *Fischer/Jeremias/Dieterich, Prüfungsrecht*, 8. Aufl., München 2022.

³ S. hierzu *Haake*, OdW 2023, 235; *dies.*, OdW 2022, 215; *dies.*, OdW 2021, 201; *dies.*, OdW 2021, 59.

Perpetuierung der Antworten von ChatGPT jedoch praktischen Schwierigkeiten. Die Abgrenzung zwischen unerlaubter und erlaubter Hilfe erfolge in gewissem Umfang auch anhand der Bewertungsrelevanz der konkreten Leistung;⁴ Sofern die Rechtschreibung Teil der Prüfungsleistung sei, könne daher auch eine rein sprachliche KI-gestützte Überarbeitung des Textes eine relevante Täuschungshandlung darstellen.

Große Herausforderungen sahen die Referenten beim Nachweis eines unzulässigen KI-Einsatzes seitens der Prüflinge. Die Beweislast für eine Täuschungshandlung liege bei den Hochschulen. Allerdings gelte auch im Prüfungsrecht zum erleichterten Nachweis bestimmter Tatsachen der Grundsatz des Anscheinsbeweises. Dessen genaue Voraussetzungen und Folgen erläuterte *Fischer* anhand einer Betrachtung der jüngeren Rechtsprechung zur Täuschung im Rahmen von Online-Prüfungen.⁵ Nach den Grundsätzen des Anscheinsbeweises genüge es, wenn die Hochschulen einen Sachverhalt vortragen, der bei typischem Geschehensablauf nahelegt, dass der Prüfling auf unerlaubte Hilfsmittel zurückgegriffen hat. Doch trotz dieser Beweiserleichterung gestalte sich die Aufdeckung eines unzulässigen KI-Einsatzes schon heute komplex. So müsse etwa der Gefahr begegnet werden, dass Studierende für besonders gute Leistungen sanktioniert werden.

Die Schwierigkeiten der Beweisführung verdeutlichte *Dieterich* den Teilnehmerinnen und Teilnehmern der Tagung anhand eines viel besprochenen Beschlusses des VG München.⁶ Im zugrundeliegenden Fall hatte sich ein Bachelorabsolvent nach einem erfolglosen Versuch 2022 ein Jahr später ein weiteres Mal auf einen Platz in einem Masterstudiengang der TU München beworben, für den Englischkenntnisse vorausgesetzt werden. Im Bewerbungsverfahren war jeweils ein Essay in englischer Sprache anzufertigen, um die Qualifikation für den Studiengang nachzuweisen. Nach dem erfolglosen ersten Versuch reichte der Bewerber im Sommer 2023 ein Essay ein, das bei den Verantwortlichen der TU München aufgrund seiner hohen Qualität den Verdacht weckte, KI-generiert zu sein. Nach weiteren Untersuchungen schloss die TU München den Bewerber vom laufenden Bewerbungsverfahren aus. Der Bewerber stellte in der Folge einen Antrag auf Erlass einer einstweiligen Anordnung mit dem Ziel der vorläufigen Zulassung zum Studium. In dem Verfahren hatte das Gericht daraufhin im Wesentlichen zu entscheiden, ob der Schluss der Hochschule gerechtfertigt war, dass der Bewerber das Essay unter unzulässigem Einsatz künstlicher Intelligenz angefertigt hat.

Die TU München trug zahlreiche Umstände vor, die in der Gesamtschau den unerlaubten KI-Einsatz beweisen sollten. Ausgangspunkt war die Überprüfung des Essays mit einer Software, die ergab, dass 45 % des Textes mit sehr hoher Wahrscheinlichkeit von künstlicher Intelligenz verfasst worden seien. Daraufhin hatte die Universität den Text von zwei Gutachtern beurteilen und prüfen lassen, welche diesem eine auffallende Prägnanz und Struktur konstatierten. Eine vergleichbare Qualität sei unter Bachelorabsolventen äußerst ungewöhnlich und werde selbst von erfahrenen Wissenschaftlern nicht immer erreicht. Ferner unterscheide sich das Essay augenfällig von dem im Vorjahr durch den Bewerber vorgelegten. Schließlich habe ein „Selbstversuch“ der Hochschule unter Verwendung von ChatGPT ein Resultat ergeben, das auffällige Ähnlichkeit zu dem vom Bewerber eingereichten Essay besitze.

Diese und weitere Umstände ließen nach der Auffassung des Gerichts unter Anwendung des Anscheinsbeweises darauf schließen, dass das Essay mit unerlaubter KI-Hilfe erstellt worden war. *Dieterich* kritisierte, dass ohne nähere Erläuterung die Regelungen des Anscheinsbeweises herangezogen wurden. Im Bereich des KI-Einsatzes gebe es bisher wohl eher kein allgemeines Erfahrungswissen, so dass kaum von einem typischen Geschehensablauf die Rede sein könne. Legte man aber die Grundsätze der in § 108 Abs. 1 Satz 1 VwGO vorgesehenen freien Beweiswürdigung im verwaltungsgerichtlichen Verfahren an, seien die Schlussfolgerungen des

⁴ Vertiefend wurde auf *Birnbaum*, NVwZ 2023, 1127 (1128) hingewiesen.

⁵ VG Berlin, Urt. v. 6.2.2023 – 12 K 52/22; VG Köln, Urt. v. 6.12.2022 – 6 K 1428/22; LG Frankfurt, Urt. 5.12.2022 – 2-01 S 89/22; Sächs. OVG, Beschl. v. 16.2.2022 – 2 B 274/21.

⁶ VG München NJW 2024, 1052 ff.; *Birnbaum*, NVwZ 2024, 607 f.; *Rachut*, NJW 2024, 1057; *Braegelmann*, RD 2024, 188 ff.

Gerichts in seinen Augen aber gut nachvollziehbar, wobei Einzelheiten der Überzeugungsbildung – etwa die Umstände des „Selbstversuchs“ mit ChatGPT – etwas unklar blieben.

Vor dem Hintergrund der erläuterten Beweisschwierigkeiten warfen *Fischer* und *Dieterich* die Frage auf, inwiefern sich Prüfungsarten und -formate zukünftig verändern müssen oder sollten. In jedem Fall haben die Hochschulen eine Weichenstellung vorzunehmen. Soll der Status Quo gewahrt werden und der KI-Einsatz in Prüfungen verboten bleiben? In diesem Fall müssen sich die Hochschulen mit der fortschreitenden Entwicklung der Technologie auseinandersetzen. Der Beschluss des VG München zeige, dass sich der unerlaubte KI-Einsatz derzeit mit einigem Aufwand noch erkennen lässt. Möglicherweise seien KI-generierte Texte aber bald zumindest für Menschen nicht mehr oder nur mit erheblichem Aufwand als solche identifizierbar. Sofern alternativ auf eine Software zurückgegriffen werden soll, stelle sich bereits die Frage, ab welcher Wahrscheinlichkeitsgrenze ein Bewerber von dem Verfahren ausgeschlossen wird. *Fischer* schlug deshalb vor, dem Einsatz künstlicher Intelligenz angepasste Aufgabenstellungen entgegenzusetzen. So könne von den Studierenden verlangt werden, ihre Lösungen auf Materialien und Ergebnissen der Vorlesungen aufzubauen. Darüber hinaus können kreative Lösungsansätze die bisher oftmals übliche Wiedergabe von Wissen als wichtigste – und leicht durch KI zu generierende – Bewertungsgrundlage ersetzen. Ein anderer Ansatz sei es, das Risiko einer Entdeckung für die Studierenden zu erhöhen. Dazu können bei entsprechender normativer Grundlage etwa eidesstattliche Versicherungen über die selbstständige Bearbeitung einer Prüfung verlangt werden, so dass der unzulässige KI-Einsatz des Prüflings strafrechtliche Konsequenzen hätte.

Ein anderer Ansatz sei es, den KI-Einsatz zuzulassen und die Prüfungsformate anzupassen. Dazu könne etwa ein Fokus auf mündliche Prüfungen und Präsenzklausuren gelegt oder eine Kombination aus Hausarbeit bzw. Online-Klausur und mündlicher Prüfung etabliert werden. Derartige Prüfungsformate seien allerdings zunächst in der Prüfungsordnung zu regeln.

2. KI-Einsatz seitens des Prüfers

Zum Abschluss des Vortrags änderte *Fischer* die Perspektive und setzte sich mit der Frage auseinander, ob der Prüfer KI-Systeme einsetzen dürfe. Die Erstellung einer weiterhin vom Prüfer letztverantworteten Prüfungsaufgabe unter Zuhilfenahme von KI sei prüfungsrechtlich nicht zu beanstanden. Problematisch sei hingegen die Bewertung einer Prüfungsleistung durch ein KI-System. Das Prüfungsrecht fordere, dass der Prüfer die Arbeit höchstpersönlich und eigenständig bewertet. Er dürfe zwar auf Assistenzen zurückgreifen. Die letzte Entscheidung dürfe er aber nicht aus der Hand geben. Darauf basierend vertrat *Dieterich* die Auffassung, dass der assistierende Einsatz künstlicher Intelligenz durch den Prüfer nicht anders zu behandeln sei als der in gewissem Umfang zulässige Rückgriff auf Korrekturassistenten. In der anschließenden Diskussion widersprach *Schwartzmann* dieser Ansicht: Die Zulässigkeit der Korrekturassistenten entspringe einem anderen Regulierungsbereich als die Frage der Zulässigkeit des KI-Einsatzes. Letztere müsse immer, aber auch ausschließlich an der Fähigkeit künstlicher Intelligenz zur Autonomie gemessen werden. Dann aber verbiete es sich, die Wertungen zur Zulässigkeit menschlicher Assistenzen auf die künstliche Intelligenz zu übertragen.

II. Voraussetzungen einer Implementierung an Hochschulen

Neben die Frage der prüfungsrechtlichen Zulässigkeit des KI-Einsatzes tritt die Frage der rechtskonformen Gestaltung und Implementierung von KI-Systemen. Letztere wird künftig abschließend von der europäischen KI-Verordnung⁷ beantwortet, welche am 1 August 2024 in Kraft getreten ist.⁸ Im Anschluss an den Vortrag *Fischers* und *Dieterichs* stellte *Schwartzmann* den

⁷ Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz), im Folgenden: KI-VO).

⁸ Im Sommer erscheint unter Mitherausgeberschaft *Schwartzmanns* der Praxisleitfaden zur KI-Verordnung *Schwartzmann/Keber/Zenner* (Hrsg.), KI-Verordnung. Leitfaden für die Praxis.

Teilnehmerinnen und Teilnehmern der Tagung das neue Regelwerk vor und beleuchtete die Konsequenzen für den Hochschulbetrieb.⁹

1. Die Regulierung von KI-Systemen durch die KI-VO

Grundsätzlich sei die KI-VO ausschließlich auf KI-Systeme anwendbar. Deren entscheidendes Merkmal sei ihre Autonomie. Ähnlich wie ein Tier könne ein KI-System zwar vom Menschen trainiert werden. Dies gewährleiste aber keinen abschließenden Schutz vor plötzlichen Ausbrüchen aus dem menschengemachten Rahmen. Die Gefahren, die sich aus dieser besonderen Fähigkeit der KI-Systeme ergeben, adressiere die KI-VO nach einem risikobasierten Ansatz. Sie ordne jedes System einer von drei Risikostufen zu. Hiernach seien die meisten Systeme unproblematisch und unterliegen deshalb keiner spezifischen Regulierung nach der KI-VO. Das gelte etwa, wenn ChatGPT eingesetzt wird, um ein Liebesgedicht oder ein Kochrezept zu schreiben. Von anderen KI-Systemen gehe nach der Einschätzung des europäischen Gesetzgebers hingegen ein hohes Risiko aus.¹⁰ Hierunter fallen etwa Systeme, die über den Hochschulzugang entscheiden oder Prüfungen bewerten. Einige Systeme seien schließlich gänzlich verboten, wie beispielsweise Systeme zur Bewertung natürlicher Personen auf Grundlage ihres sozialen Verhaltens. Eine Schwierigkeit machte *Schwartzmann* bei der Abgrenzung zwischen hochriskanten und verbotenen KI-Systemen aus. Ermittelt eine Hochschule etwa per KI, ob ihre Studierenden unter Prüfungsangst leiden und zieht daraus Schlüsse für die Prüfungspraxis, könne es schwerfallen, zwischen einem verbotenen Social Scoring und einem unter strengen Voraussetzungen erlaubten Einsatz im Hochschulkontext zu unterscheiden.

Um den „wilden“ Einsatz von KI-Systemen wie ChatGPT durch private Accounts zu unterbinden, empfahl *Schwartzmann* entsprechende Produkte zu lizenzieren. Nur so können die Chancengleichheit gewahrt und lizenzrechtliche Probleme vermieden werden. Mit der Lizenzierung werde die Hochschule aber zugleich Betreiberin eines KI-Systems im Sinne der KI-VO. Sie sei daher unter anderem verpflichtet, sicherzustellen, dass ihr Personal über hinreichende KI-Kompetenz verfügt. Außerdem müsse sie KI-generierte Inhalte als solche kennzeichnen. Falls das lizenzierte KI-System nach der KI-VO hochriskant sei, müsse die Hochschule schließlich vor dessen Inbetriebnahme eine Grundrechte-Folgenabschätzung vornehmen. Dabei sei zu prüfen, welche Auswirkungen die Verwendung des Systems auf die Grundrechte der betroffenen Personen haben kann.

Wesentlich weiter als die genannten Betreiberpflichten gingen die Pflichten der Anbieter von Hochrisiko-KI-Systemen. Grundsätzlich fallen darunter zwar nur Systementwickler wie OpenAI oder AlephAlpha. Unter Umständen können aber auch Betreiber in die Anbieterstellung geraten. Lizenziert eine Hochschule etwa ein KI-System, das im Zugangsverfahren assistieren soll, und versieht dieses mit dem eigenen Logo, habe sie in der Folge die umfangreichen Pflichten der Hochrisiko-KI-Anbieter zu erfüllen. So habe sie etwa sicherzustellen, dass der Betrieb des Systems von einer menschlichen Aufsicht überwacht werden kann. Ebenso werde die Hochschule von der KI-VO als Anbieterin eines Hochrisiko-KI-Systems behandelt, wenn sie ein lizenziertes Hochrisiko-KI-System wesentlich verändert, etwa durch technische Anpassung einer Bewertungs-KI an besondere Prüfungsformate.

Darüber hinaus seien Betreiber und sonstige Dritte nach der KI-VO als Anbieter eines Hochrisiko-KI-Systems zu behandeln, wenn sie die Zweckbestimmung eines KI-Systems so verändern, dass das betreffende KI-System nach der KI-VO als hochriskant zu klassifizieren ist. Diese Regelung gelte auch für KI-Systeme mit originär allgemeinen Verwendungszweck. ChatGPT diene etwa der Generierung von Texten und sei damit nicht zweckgebunden. Es gelte daher im Grundsatz nicht als hochriskant. Wenn allerdings ein Prüfer das KI-System einsetze, um eine Hausarbeit zu bewerten, nutze er das System zu einem hochriskanten Zweck. In der

⁹ Zum Ganzen wurde vertiefend auf folgende Quellen hingewiesen: *Schwartzmann*, F.A.Z. v. 24.4.2024, S. N4; *Schwartzmann*, Forschung und Lehre 2024, 348 (349); *Benedikt/Köhler/Schwartzmann/Wünschelbaum* F.A.Z v. 18.3.2024, 18; F.A.Z. Einspruch Podcast, Folge 292.

¹⁰ Vgl. die nicht abschließende Aufzählung in Anhang III KI-VO.

Folge seien die Anforderungen an Hochrisiko-KI-Systeme zu erfüllen. Wolle die Hochschule einer Verantwortlichkeit für das Verhalten des Prüfers entgehen, müsse sie ihrem Personal den Einsatz des Systems zu hochriskanten Zwecken untersagen. Handle ein Mitarbeiter entgegen dieser Anweisung, so befinde er sich im Exzess und habe für die Folgen persönlich einzustehen.

2. Prüfungsbewertungen von ChatGPT?

Die praktischen Schwierigkeiten beim Einsatz von ChatGPT zur Bewertung von Prüfungsarbeiten erläuterte *Schwartmann* anhand eines Selbstversuchs, den er gemeinsam mit einer Mitarbeiterin und ehemaligen Studentin durchgeführt hatte: In dem Versuch wurde das KI-System mit den Bewertungsvorgaben der geltenden Prüfungsordnung und der Masterarbeit der Studentin gefüttert. Seine Aufgabe lautete, die Arbeit anhand der verfügbaren Informationen zu benoten. Tatsächlich bewertete ChatGPT die Arbeit mit 1,3 – der gleichen Note, mit der auch die Prüfer die Arbeit bewertet hatten. Daraufhin änderte *Schwartmann* das Thema der Arbeit. Statt um den Einsatz von KI im Beschäftigtendatenschutz sollte die Arbeit nun den Einsatz von KI im Baurecht behandeln. Jeder Prüfer müsse in dieser Konstellation erkennen, dass die eingereichte Masterarbeit die vorgegebene Aufgabenstellung nicht erfüllt. Thema verfehlt. Nicht so die KI: ChatGPT gab auch auf die veränderten Gegebenheiten hin an, dass die Arbeit mit der Note 1,3 zu bewerten sei.

Schwartmann wies darauf hin, dass das Ergebnis seines Selbstversuchs nicht auf einen Fehler der KI zurückzuführen sei, sondern auf einen menschlichen Fehler in der Anwendung. Nur durch genaue menschliche Vorgaben und einen bedachten Einsatz könne die KI einen echten Mehrwert liefern. Es gehe daher eher um Verbesserung der Qualität als um Arbeitserleichterung per Zauberhand. Um einen verantwortungsbewussten Gebrauch der KI-Systeme sicherzustellen, müsse die in der KI-VO vorgesehene Stopptaste für Hochrisiko-KI-Systeme als Starttaste verstanden werden: Anstatt die KI abzuschalten, wenn es womöglich bereits zu spät ist, solle der Mensch sich eigene Gedanken machen und die KI sodann bewusst zur Überprüfung dieser eigenen Gedanken starten.

III. Prüfungen in Zeiten künstlicher Intelligenz

In der anschließenden Fragerunde wurden zunächst Verfahren diskutiert, um einen verbotenen Einsatz von KI-Anwendungen in Prüfungssituationen zu verhindern. Vorgeschlagen aus dem Kreis der Zuhörer wurde unter anderem, dass bei Hausarbeiten Studierende ausgelost werden könnten, die zusätzlich zur schriftlichen Leistung einen mündlichen Vortrag halten müssen. Der mündliche Vortrag könne außerhalb der Bewertung der Hausarbeit zu halten sein, oder aber Bestandteil der Bewertung werden. So solle generalpräventiv sichergestellt werden, dass die Studierenden die geforderte Leistung eigenständig erbringen bzw. die fehlende Eigenständigkeit bei der Erstellung der schriftlichen Leistung aufgedeckt werden. *Fischer* verwies auf die Parallelen zwischen Täuschungsversuchen mit KI-Anwendungen in digitalen Prüfungen: Bei Letzteren seien bereits Fälle bekannt, in denen Prüfer nach einer schriftlichen Leistung in einem kurzen Gespräch abgefragt hätten, ob der jeweilige Prüfling den Stoff und die Entstehung der Arbeit wiedergeben kann. *Schwartmann* stellte in diesem Kontext einen Ansatz vor, den er selbst gerade mit den Studierenden des von der TH Köln angebotenen Masterstudiengangs „Medienrecht und Medienwirtschaft“ erprobt: Die Studierenden erarbeiten dabei ein Thema unter Einsatz künstlicher Intelligenz. Im Anschluss präsentieren sie die KI-generierten Ergebnisse in einem mündlichen Vortrag und setzen sich kritisch mit diesen auseinander.

Fischer wies darauf hin, dass, falls das Gespräch Teil der zu erbringenden Leistung sein solle, dies in der Prüfungsordnung festzuhalten sei. Eine Auslosung einiger weniger Studierender würde in diesem Fall – anders als bei Stichproben zur Aufdeckung von Täuschungsversuchen – dann auch einen Verstoß gegen die Chancengleichheit darstellen und verbiete sich deshalb. Doch auch wenn das Gespräch nur zur Überprüfung eines Täuschungsverdachts vorgesehen sei, empfehle sich eine Regelung in der Prüfungsordnung zur Mitwirkungspflicht des Prüflings bei der Aufklärung des Sachverhalts. *Dieterich* ergänzte, dass sich im Übrigen auch § 26 VwVfG für Grund und Grenze der Mitwirkungspflicht bemühen lässt. Hiernach sollen alle Beteiligten eines Verwaltungsverfahrens bei der Ermittlung des Sachverhalts mitwirken. Eine Pflicht zum

persönlichen Erscheinen oder zur Aussage bestehe nach § 26 Abs. 2 VwVfG aber nur, soweit sie durch Rechtsvorschrift besonders vorgesehen ist.

Auf reges Interesse stießen auch einige Regelungen, die nun von der KI-VO eingeführt werden. So kam die Frage auf, ob die Transparenzpflicht beim Einsatz textgenerierender KI-Systeme auch für Studierende gelte, die ChatGPT in ihren Prüfungen einsetzen. *Schwartmann* stellte klar, dass die Studierenden keine Anbieter im Sinne der KI-VO sind und damit nicht von der Regelung zur Transparenz adressiert werden. Ebenso gerieten sie nicht über eine Zweckänderung in die Anbieterrolle, da der Einsatz von KI-Systemen durch Studierende von der KI-VO nicht als hochriskant klassifiziert werde. Ferner wurde gefragt, ob die Verpflichtung zur Durchführung einer Grundrechte-Folgenabschätzung bereits greife, wenn ein Prüfer spontan und entgegen der Weisung der Hochschule ChatGPT zu Bewertungszwecken einsetze. *Schwartmann* betonte, dass die Verpflichtung zur Durchführung einer Grundrechte-Folgenabschätzung ausschließlich für die Betreiber von Hochrisiko-KI-Systemen gilt. ChatGPT sei als KI-System mit allgemeinem Verwendungszweck zwar nicht grundsätzlich hochriskant. Sein Einsatz zu hochriskanten Zwecken löse aber die spezifische Hochrisiko-Regulierung nach der KI-VO aus. Der Prüfer werde in dem Fallbeispiel zum Anbieter eines Hochrisiko-KI-Systems und müsse daher umfangreiche Pflichten erfüllen. Hinzu trete die an Betreiber adressierte Grundrechte-Folgenabschätzung, die aber in Anbetracht der Anbieterpflichten einen zu vernachlässigenden Mehraufwand für den Prüfer bedeute.

Schließlich lieferte *Dieterich* noch einen Überblick zum aktuellen Stand in der Diskussion um die Zulässigkeit sog. Proctoring-Maßnahmen. Dabei handelt es sich um Überwachungsmaßnahmen zur Prävention von Täuschungen in digitalen Prüfungen, etwa per Videoaufsicht. Bereits im Rahmen der ersten Tagung zu Digitalen Prüfungen war diese Frage kontrovers diskutiert worden.¹¹ Das OVG NRW hatte im Frühjahr 2021 im Eilrechtsschutzverfahren entschieden, dass eine Videoaufsicht während digitaler Prüfungen prüfungsrechtlich geboten und datenschutzrechtlich grundsätzlich zulässig sei.¹² *Dieterich* konstatierte, dass mit der damaligen Begründung des Gerichts eine Videoaufsicht mittlerweile wohl nicht mehr zu rechtfertigen sei. In seiner Entscheidung während der Corona-Pandemie habe das Gericht noch wesentlich darauf abgestellt, dass es ein öffentliches Interesse am Fortschreiten der Ausbildung gebe und ein Anspruch auf die Durchführung von Prüfungen bestehe, was jedenfalls die Möglichkeit einer Rechtfertigung der Datenverarbeitung über Art. 6 Abs. 1 UAbs. 1 Buchst. e) DS-GVO eröffne. Da Präsenzprüfungen derzeit aber möglich seien, entfalle diese Möglichkeit. Stattdessen müsse eine Videoaufsicht nun auf die freiwillige Einwilligung der Prüflinge gestützt werden (vgl. Art. 6 Abs. 1 UAbs. 1 Buchst. e) DS-GVO), was für die Prüfungsbehörden aufgrund der Widerrufsmöglichkeit mit rechtlicher Unsicherheit verbunden sei (vgl. Art. 7 Abs. 3 DS-GVO).¹³

IV. Resümee und Ausblick

Der Einsatz künstlicher Intelligenz in Prüfungssituationen stellt das materielle Prüfungsrecht vor keine allzu großen Herausforderungen. Wenn ein Prüfling ein KI-System ohne Kenntlichmachung einsetzt, um eine bewertungsrelevante Leistung zu erbringen, nimmt er eine sanktionierbare Täuschungshandlung vor. Weitaus diffiziler gestaltet sich der Nachweis einer solchen Handlung. Trotz der Beweiserleichterungen durch die Anwendung der Grundsätze des Anscheinsbeweises ist die Feststellung eines unzulässigen KI-Einsatzes schon heute mit erheblichem Aufwand verbunden, der voraussichtlich mit wachsender Komplexität und Leistungsfähigkeit der Systeme weiter anwachsen wird. Vor diesem Hintergrund sollten Hochschulen neue Prüfungsformate in den Blick nehmen, die einen stärkeren Fokus auf eine Leistungserbringung in Präsenz oder in mündlichen Prüfungen legen.

Zunächst aber gilt es, KI-Systeme rechtskonform in den Hochschulalltag zu integrieren. Dabei sind die Regelungen der KI-VO zu beachten, die im Sommer in Kraft treten wird. Insbesondere

¹¹ *Haake*, OdW 2023, 235 (60 und 63 f.).

¹² OVG NRW NJW 2021, 1414.

¹³ Vertiefend wurde auf *Dieterich* NVwZ 2021, 511 (515 ff.) hingewiesen.

der Einsatz von KI-Systemen zu Bewertungszwecken oder im Zulassungsverfahren wird von der KI-VO als hochriskant klassifiziert. Um die pflichtenreiche Anbieterrolle zu vermeiden, sollten Hochschulen KI-Systeme für klar definierte Aufgabenfelder lizenzieren und ihr Personal anweisen, die Systeme ausschließlich zu den vorgegebenen Zwecken einzusetzen. Sofern sie durch diese Maßnahmen in der Betreiberrolle verbleiben, müssen die Hochschulen vor allem für ein ausreichendes Maß an KI-Kompetenz sorgen und eine Grundrechte-Folgenabschätzung durchführen.

Die eingangs zitierte Bitkom-Studie zeigt, dass KI-Systeme an deutschen Hochschulen längst der Realität entsprechen. Mit der Tagung wurde nun auch der Grundstein für deren verantwortungsbewussten und rechtssicheren Einsatz gelegt. In seinen Schlussworten wies *Epping* darauf hin, dass die prüfungsrechtliche Behandlung künstlicher Intelligenz damit allerdings nicht abgeschlossen sei. Bei entsprechendem Interesse der Teilnehmenden beabsichtige der Verein zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. deshalb, im kommenden Jahr eine Anschlussveranstaltung zu organisieren, in der dann weitere Erkenntnisse zum KI-Einsatz an Hochschulen mit Interessierten geteilt werden sollen.

Moritz Köhler ist wissenschaftlicher Mitarbeiter der Kölner Forschungsstelle für Medienrecht und Doktorand bei Prof. Dr. Rolf Schwartmann.