

Prüfung und Täuschung – Vom Spickzettel bis zum Einsatz Künstlicher Intelligenz

Bericht über die Tagung des Vereins zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. am 7. Oktober 2025

Von Dr. Rico David Neugärtner, LL.M. (Cornell), Berlin

Die Täuschung ist ein Klassiker des Prüfungsrechts. Klassiker können aber immer wieder neu gelesen werden. Die Tagung „Prüfung und Täuschung“ des Vereins zur Förderung des deutschen und internationalen Wissenschaftsrechts e.V. (VFDIW) am 7. Oktober 2025 unternahm einen Versuch der Lektüre von (vermeintlich) Bekanntem in neuem Licht. Dabei war der Untertitel „Vom Spickzettel bis zum Einsatz Künstlicher Intelligenz“ Programm: Er beschreibt den medialen Bogen aus der romantisch-analogen Welt des Handgekritzelten in die faszinierend-beängstigende Zukunftsgegenwart der Künstlichen Intelligenz (KI). Zwei der Hauptaussagen, die von der Veranstaltung nachhallen, sind diese: Die allgemeine Prüfungsrechtsdogmatik ist gut gewappnet, trotz medialer und technischer Innovationen Orientierung und stabile Lösungsangebote zu liefern – nicht alles muss neu erfunden werden –; es ist aber zumindest zweckmäßig, an mancher Stelle klug zu gestalten, zum Beispiel beim behutsamen Fortentwickeln von Aufgabenformaten, Prüfungsordnungen und Infrastrukturen.

Die Online-Tagung vom 7. Oktober 2025 stellte sich in eine Reihe von Veranstaltungen des VFDIW zu Neuerungen im Prüfungsrecht in Folge der galoppierenden Digitalisierung.¹ Zur Eröffnung blickte Prof. Dr. Volker Epping (Präsident der Leibniz Universität Hannover und Vorstandsmitglied des VFDIW) nach Begrüßung der – beinahe 200 – digital zugeschalteten Teilnehmerinnen und Teilnehmer knapp auf die Geschichte der Veranstaltungsreihe zurück, um sodann die Dozenten des Tages zu präsentieren: Edgar Fischer, Vorsitzender Richter am Verwaltungsgericht Berlin, und Dr. Peter Dieterich, LL.M. (Chicago), Richter am Verwaltungsgericht Berlin.²

Diese sollten in drei Blöcken jeweils im Rahmen von Impulsreferaten Allgemeines zur Täuschung im Prüfungsrecht (dazu I.), dann Besonderheiten von Täuschungen bei Online-Prüfungen (dazu II.) und schließlich Spezifisches zum Einsatz von künstlicher Intelligenz bei Prüfungen (dazu III.) skizzieren, um sodann jeweils mit den Teilnehmerinnen und Teilnehmern

¹ Berichte zu früheren Veranstaltungen: Haake, OdW 2021, 59 ff.; dies., OdW 2021, 201 ff.; dies., OdW 2022, 215 ff.; dies., OdW 2023, 235 ff.; Köhler, OdW 2024, 339 ff.; Alt, OdW 2025, 201 ff.

² Fischer und Dr. Dieterich sind neben Dr. Christoph Jeremias Autoren von Fischer/Jeremias/Dieterich, Prüfungsrecht, 9. Aufl. 2026, im Erscheinen.

zu Einzelfragen ins Gespräch zu kommen. Die drei Diskussionsrunden moderierte Prof. *Ulf Pallme König* (Universitätskanzler a.D., Ehrenvorsitzender des VFDIW). Im Folgenden wird von den wesentlichen Inhalten der drei Impulsreferate sowie den substanzialen Gegenständen der anschließenden Diskussionsrunden in verzahnter Weise berichtet; das heißt: über die Diskussionsgehalte wird an thematisch jeweils einschlägiger Stelle informiert.

I. Allgemeines zur Täuschung im Prüfungsrecht

1. „Täuschung“: Vielgestaltigkeit der relevanten Handlungen

In das erste Impulsreferat stieg *Fischer* mit einer Vorstellung des objektiven Tatbestands der „Täuschung“ ein: Dieser liege in der Vorspiegung einer selbstständigen und regulären Prüfungsleistung, obwohl in Wahrheit unerlaubte oder nicht offen gelegte Hilfen genutzt worden seien. Dabei hob *Fischer* die Technologieoffenheit der allgemeinen Dogmatik hervor: Die allgemeine Begriffsbestimmung erfasse ohne Weiteres etwa auch den nicht erlaubten oder die – im Fall einer Erlaubnis – nicht hinreichend offengelegte Nutzung von KI, etwa beim Verfassen einer Hausarbeit. Dieser Einsatz sei nicht anders zu beurteilen, als wenn ein anderer die Aufgabe für den Prüfling löst.

Fischer führte weiter aus, dass grundsätzlich bereits der Versuch einer Täuschung genüge, um die entsprechenden Rechtsfolgen nach der Prüfungsordnung auszulösen. Abzugrenzen sei der Täuschungsversuch von einer bloßen Vorbereitungshandlung, etwa wenn der Prüfling zu Hause einen Spickzettel für eine Aufsichtsarbeit verfasse. Mit Beginn der Aufsichtsarbeit werde die Schwelle zum Täuschungsversuch überschritten, wenn der Prüfling den Spickzettel bei sich führe.

In der späteren Diskussion stellte eine Teilnehmerin die Frage, wie es sich verhalte, wenn bei einer Klausureinsicht – also nach Abgabe der Prüfungsleistung – der Prüfungsgegenstand durch das Einschmuggeln einer angeblich bei der Korrektur übersehenen Seite nachträglich manipuliert werde. Sie äußerte Bedenken, ob die konkret einschlägige Prüfungsordnung es zulasse, diesen Sachverhalt als Täuschung zu erfassen, da diese darauf abstelle, dass das Verhalten darauf abziele, „das Ergebnis der Prüfungsleistung zu beeinflussen“, die Bewertung aber schon abgeschlossen sei. *Fischer* antwortete, es sei richtig, sorgfältig mit der konkreten Prüfungsordnung zu arbeiten, hatte aber wenig Bedenken, dass eine Täuschung auch bei dieser tatbestandlichen Ausgestaltung letztlich zu bejahen sein könnte, da weiterhin das Ergebnis der Prüfung beeinflusst werden sollte.

2. *Subjektiver Tatbestand, vor allem Anscheinsbeweis und richterliche Überzeugungsbildung*

Subjektiv, so *Fischer* weiter, sei zur Annahme einer Täuschung zumindest bedingter Vorsatz erforderlich, für welchen ein Inkaufnehmen etwa des Beisichführens eines Smartphones bei der Aufsichtsarbeit genüge. Die Prüfungsbehörde treffe auch für den subjektiven Tatbestand die Darlegungs- und Beweislast.

Als Beweiserleichterung zugunsten der Prüfungsbehörde könne im Einzelfall der Anscheinsbeweis in Betracht kommen. Dieser greife ein, wenn im Einzelfall ein typischer Geschehensablauf vorliege, der nach der Lebenserfahrung auf eine bestimmte Ursache hinweise. Nachzuweisen habe die darlegungs- und beweisbelastete Seite dann nur den Ausgangstatbestand, also beispielweise die Tatsache, dass die Prüfungsleistungen, welche von zwei Studierenden bei einer Online-Klausur eingereicht worden seien, in zahlreichen Details (etwa im Hinblick auf Formulierungen, die man auf anders hätte fassen können, auf typische Fehler, auf die man nicht ohne Weiteres kommt, auf übereinstimmende Rechtschreib- oder Formatierungsfehler) übereinstimmen. Aufgrund der feststehenden Tatsachen dränge sich dann bei verständiger Würdigung der Schluss auf, dass der Prüfungsteilnehmer getäuscht habe und ein abweichen-der Geschehensablauf nicht ernsthaft in Betracht komme. Ein Gegenbeweis bleibe dem Prüfling möglich.

Dr. *Dieterich* hob die Grenzen des Anscheinsbeweises anhand eines Judikats des Oberverwaltungsgerichts Niedersachsen³ hervor: In überzeugender Weise habe das Gericht herausgearbeitet, dass die Voraussetzungen für den Anscheinsbeweis nicht vorschnell angekommen werden dürften. So reiche etwa nicht als Ausgangstatbestand, wenn die bei einer juristischen Klausur vorgelegte Lösung der Lösungsskizze in einzelnen Punkten ähnle, um darauf schließen zu können, der Prüfling müsse bei lebensnaher Betrachtung die Lösungsskizze bei Leistungserbringung gekannt haben. Denn schematische Lösungswege seien gerade bei juristischen Prüfungen nicht unüblich. Das Gericht habe sodann geprüft – und trotz gewichtiger Anhaltspunkte etwas überraschend letztlich verneint –, ob die Täuschung ohne die Erleichterung des Anscheinsbeweises im Rahmen der richterlichen Überzeugungsbildung nach den allgemeinen Beweisregeln (vgl. § 108 Abs. 1 Satz 1 VwGO) gewonnen werden könne.

³ Urteil vom 30. April 2024 – 2 LB 69/18 – juris; vgl. nachgehend BVerwG, Beschluss vom 20. Januar 2025 – 6 B 20/24 – juris.

3. Umgang mit Täuschungsverdacht vor Ort; Täuschungsprävention

Einen starken Praxisbezug wiesen die nachfolgenden Empfehlungen von *Fischer* zum prüfungsbehördlichen Umgang mit einem Täuschungsverdacht während der laufenden Prüfung auf. Aufsichtspersonal solle selbstbewusst auftreten, Beweise sichern und die beobachteten Vorgänge samt den Reaktionen des Prüflings im Protokoll festhalten. Grundsätzlich sei es sinnvoll, dem Prüfling das Fortsetzen der Leistungserbringung zu ermöglichen, um Folgeauseinandersetzungen zu vermeiden. Unter Umständen könnten kurzfristig Maßnahmen wie das Umsetzen eines täuschungsverdächtigen Prüflings zweckmäßig sein. In diesem Zusammenhang ging *Fischer* auf Judikate des Verwaltungsgerichts München⁴ ein, denen ein Sachverhalt zugrunde gelegen habe, bei dem die Prüfungsaufsicht wegen Täuschungsverdachts einen Prüfling im Saal weiter nach vorn versetzt und abgetastet habe.

Zu vermeiden seien voreilige und endgültige Maßnahmen. Das wurde in der späteren Diskussion anschaulich, als eine Teilnehmerin schilderte, zum Teil sei Aufsichtspersonal nicht hinreichend informiert, beispielsweise welche Kennzeichnungen in Gesetzestexten bei juristischen Aufsichtsarbeiten erlaubt seien. *Fischer* empfahl, Vorsorge zu treffen, damit stets ein schneller Kontakt zu einer fachlich informierten Person gewährleistet sei.

Eine weitere Teilnehmerin kam auf Formen der Täuschungsprävention zu sprechen: Sie fragte nach der rechtlichen Zulässigkeit des Einsatzes von Handydetektoren („Sniffen“) vor Prüfungsbeginn. Die Dozenten stellten unter Verweis auf die prüfungsbehördliche Organisationskompetenz klar, dass es der Prüfungsaufsicht grundsätzlich gestattet sei, Hilfsmittel zur Gewährleistung des ordnungsgemäßen Prüfungsverfahrens einzusetzen. Bei grundrechtsintensiven Kontrollen könnte eine Regelung in der Prüfungsordnung zweckdienlich und gegebenenfalls sogar rechtlich geboten sein. Dr. *Dieterich* wies darauf hin, dass sich in Extremsfällen materielle verfassungsrechtliche Grenzen, insbesondere aus dem Allgemeinen Persönlichkeitsrecht, ergeben könnten, etwa wenn ein Studierender einen Spickzettel herunterschlänge; dann sei die Aufsicht gehindert, diesen zu zwingen, den Zettel preiszugeben, könne aber das Verhalten dokumentieren und der Prüfungsbehörde damit ermöglichen, hieraus Rückschlüsse zu ziehen.

⁴ Beschluss vom 30. April 2020 – M 3 E 20.1243 – juris Rn. 37; ferner Urteil vom 25. Oktober 2022 – M 3 K 20.650 – juris.

4. Abgestufte Sanktionen; schwere Täuschungen

Im Impulsreferat selbst ging *Fischer* als nächstes auf die gängige Grundsystematik von Prüfungsordnungen hinsichtlich der Sanktionen bei festgestellten Täuschungen ein. Typischerweise seien die Sanktionen nach Schwere der Täuschung ausdifferenziert. Sollte auf Rechtsfolgenebene Ermessen eingeräumt werden, müsse die Prüfungsbehörde ihr Ermessen auch ausüben und dies zum Ausdruck bringen. Im Übrigen würden Prüfungsordnungen häufig für den Fall „schwerer Täuschungen“ einschneidendere Sanktionen bis hin zum endgültigen Nichtbestehen vorsehen. *Fischer* referierte zu unterschiedlichen Tendenzen in der Rechtsprechung hinsichtlich der Frage, ob die Nutzung eines Smartphones eine schwere Täuschung darstelle.⁵

Vorsicht sei geboten, wenn die Prüfungsordnung das Merkmal eines „besonders schweren Falls“ auf Tatbestandsebene mit Ermessen auf der Rechtsfolgenebene kombiniere. Hier bestehe die Versuchung, dass die Prüfungsbehörde die Subsumtion unter das genannte Tatbestandsmerkmal bereits unter umfassender Auswertung des Einzelfalls vornehme und sodann bei ihrer Ermessensausübung diese Umstände schlicht nochmals wiederhole. Dies sei rechtlich problematisch. Diese Thematik wurde in der späteren Diskussion mehrfach aufgegriffen. Nach Dr. *Dieterich* sei das ein Beispiel dafür, dass „schlanke Prüfungsordnungen“ im Zweifel praktikabler sein könnten. So sei das skizzierte Problem vermeidbar, wenn die Prüfungsordnung die Sanktionierungsstufen nicht schon auf Tatbestands-, sondern erst auf Ermessensebene ausdifferenziere. Liege nun aber eine Prüfungsordnung mit der genannten Kombination von „schwerer Täuschung“ als Tatbestandsmerkmal und Ermessen auf Rechtsfolgenseite vor, müsse die Behörde damit umgehen und sich etwa ein Argument für die Ermessensausübung aufsparen oder dort die zuvor auf der Tatbestandsebene benannten Umstände besonders qualifizieren.

5. Plagiate

Zum Abschluss des ersten Teils und zur Vorbereitung der späteren Erörterungen zu Plagiaten bei neuen Prüfungsformaten (siehe unten II. 4.) und zu KI-Plagiaten (siehe unten III.) referierte *Fischer* zu der allgemeinen Dogmatik der Täuschung durch Plagiieren. Er erläuterte

⁵ Vgl. für eine schwere Täuschung OVG Nordrhein-Westfalen, Beschluss vom 16. Februar 2021 – 6 B 1868/20 – ; dagegen etwa Bayerischer VGH, Beschluss vom 5. November 2024 – 7 ZB 24.632 – juris Rn. 10 („Das Smartphone [...] war ausgeschaltet und vom Kläger während der Prüfung nicht verwendet worden. Auf dem Smartphone waren keine prüfungsbezogenen Daten gespeichert“; gegen schwere Täuschung auch VG Gelsenkirchen, Urteil vom 15. Januar 2025 – 4 K 3460/21 – juris Rn. 52: „Denkbar ist [...], dass [der Prüfling] mithilfe seines Smartphones – ähnlich einem Spickzettel – nur einzelne Wissenslücken schließen wollte.“)

insbesondere die Formel zur Ermittlung von Plagiatsstellen, wonach „der Leser an jeder Stelle [wissen können müsse], wer zu ihm spricht“⁶, und den Ansatz des Bundesverwaltungsgerichts zur Bestimmung der Erheblichkeit von Plagiatsstellen für die (Gesamt-)Prüfungsleistung, wonach zu fragen ist, ob „[d]ie Plagiatsstellen [...] die Arbeit quantitativ, qualitativ oder in einer Gesamtschau beider Möglichkeiten prägen“⁷.

II. Täuschung bei Online-Prüfungen

1. Begriffliches, vor allem Abgrenzung von der ‚elektronischen Prüfung‘

Auf der so geschaffenen Basis des Altbewährten ordnete sodann Dr. *Dieterich* zu Beginn des zweiten Impulsvortrags das Begriffswirrsal rund um die ‚Online-Prüfung‘: Nach den Dozenten sei darunter eine Prüfung zu verstehen, welche über digitale Kommunikationssysteme wie eine Videokonferenzanlage vermittelt werde, Prüfling und Prüfer dabei also nicht zugleich in ein und demselben realen Prüfungsraum anwesend seien. Es handele sich dabei mangels Verschiebung der abgeprüften Kompetenz jedoch nicht um eine eigene Prüfungsart; vielmehr könnten alle Prüfungsarten (schriftlich, mündlich, praktisch, elektronisch) grundsätzlich entweder herkömmlich unter Anwesenden oder eben als Online-Prüfung durchgeführt werden. Und daher bejahte Dr. *Dieterich* in der späteren Diskussionsrunde auch die Frage eines Zuhörers, ob es sich bei einer dauerhaft überwachten Online-Klausur um eine „Aufsichtsarbeit“ handele.⁸

Zu betonen sei die Abgrenzung des Begriffs ‚Online-Prüfung‘ von demjenigen der elektronischen Prüfung: Letzterer bezeichne eine eigenständige Prüfungsart, bei welcher bei der Erbringung der Prüfungsleistung Daten in eine von der Prüfungsbehörde zur Verfügung gestellte EDV-Maske einzugeben seien – dies sei nicht nur online, sondern etwa auch in einem Computer-Pool der Prüfungsbehörde in Präsenz durchführbar. Bei elektronischen Prüfungen ließen sich die technischen Möglichkeiten der IT nutzen, um in spezifischer Weise Kompetenzen abzuprüfen. So beispielsweise durch den Einsatz von einzelaufgabenbezogenen Timern oder durch die Verhinderung eines Zurückklickens zur vorherigen, bereits beantworteten Aufgabe.

In der späteren Diskussion äußerte eine Teilnehmerin ihr Unbehagen: Sie fürchte, durch ein bloßes Umetikettieren – etwa wenn man statt einer Klausur ein *take-home exam* durchföhre – könnte die Chancengleichheit nicht hinreichend gewährleistet werden. Dr. *Dieterich*

⁶ Etwa bei VG Düsseldorf, Urteil vom 20. März 2014 – 15 K 2271/13 – juris Rn. 108.

⁷ BVerwG, Urteil vom 21. Juni 2017 – 6 C 3/16 – juris Rn. 44.

⁸ Beachte aber Fußn. 9.

antwortete, er verstehe das Störgefühl, aber *take-home exams* seien als – besondere, da kurze – Hausarbeiten, nicht als Aufsichtsarbeiten zu qualifizieren und könnten durchaus in sinnvoller Weise das Repertoire der Prüfungsformate ergänzen. Ob jemand in der Lage sei, in kurzer Zeit mit allen möglichen Hilfsmitteln ein Problem zu bearbeiten, könne praxisrelevante Kompetenzen betreffen. Freilich müssten letztlich die Prüfenden entscheiden, ob der Einsatz eines solchen Formats – sofern nach der Prüfungsordnung möglich – zweckmäßig sei.

2. Spezifische Rechtsgrundlagen?

Eine Folge der Einordnung von Online-Prüfungen als bloßer Modifikation der bekannten Prüfungsarten sei, dass – aus prüfungsrechtlicher Sicht – eine eigenständige Rechtsgrundlage grundsätzlich nicht erforderlich sei; anders sei dies gegebenenfalls aus datenschutzrechtlicher Sicht. Freilich legte Dr. *Dieterich* offen, dass diese Frage in der Rechtsprechung nicht einheitlich beantwortet werde.⁹ Anders sei dies – das wurde später in der Diskussion klargestellt –, wenn ein echter Regelungsauftrag (nicht nur eine klarstellende Ermächtigung) etwa im Landeshochschulgesetz enthalten ist. Und natürlich sei der Prüfungsordnungsgeber frei, eine spezifische Rechtsgrundlage etwa für *take-home exams* zu schaffen.

Eine weitere Folge des von den Dozenten vertretenen Ansatzes sei, dass die tradierten dogmatischen Grundsätze, welche etwa für eine herkömmliche mündliche Prüfung entwickelt worden seien, auch auf eine mündliche Prüfung im Format einer Online-Prüfung übertragbar seien. Freilich könnte das zu Schwierigkeiten führen, etwa wenn in einer Prüfungsordnung für mündliche Prüfungen die Hochschulöffentlichkeit bestimmt werde. Dies ließe sich entweder so auflösen, dass die technischen Voraussetzungen für die Hochschulöffentlichkeit auch bei Online-Prüfungen geschaffen würden, oder so, dass eine Ausnahme von der Hochschulöffentlichkeit ausdrücklich geregelt würde.

3. Täuschungsprävention bei Online-Prüfungen

Auch bei Online-Prüfungen – zumal angesichts deren „Täuschungsgeneigtheit“ – gebe es, so Dr. *Dieterich* weiter, eine Pflicht der Prüfungsbehörde zur Täuschungsprävention.¹⁰ Diese folge aus dem Zweck der Prüfung, ein realistisches Abbild der Leistungsfähigkeit zu geben, dem gebotenen Schutz derjenigen, die später mit den Leistungen des Prüflings in Kontakt

⁹ Wie die Dozenten etwa VG Frankfurt (Oder), Beschluss vom 11. Mai 2021 – 1 L 124/21 – juris Rn. 25; offen VG Schleswig, Beschluss vom 1. März 2022 – 12 B 10003/21 – juris Rn. 11; anders VG Freiburg, Urteil vom 15. Februar 2022 – 8 K 183/21 – juris Rn. 27, 29: „Eine Aufsichtsarbeits ist [...] von der physischen Anwesenheit einer Aufsichtsperson geprägt.“

¹⁰ Vgl. auch VG Frankfurt (Oder), Beschluss vom 11. Mai 2021 – 1 L 124/21 – juris.

kommen werden, der Chancengleichheit und der relativen Bewertungsebene. Freilich gelte der Vorbehalt des (technisch, personell und finanziell) Machbaren. Man müsse sich vor Augen führen, dass auch bei Präsenzprüfungen Täuschungen nicht mit Sicherheit zu vermeiden seien.

Dr. *Dieterich* skizzierte „sanfte“ und „harte“ Präventionsmaßnahmen. Sanfte – das heißt: datenschutz- und persönlichkeitsrechtlich weitgehend unproblematische – Steuerung sei beispielsweise denkbar durch das Schaffen eines hohen Zeitdrucks, das Einfordern einer handschriftlichen Abfassung zur Vermeidung von *Copy-and-Paste* oder das Umstellen auf Transferaufgaben. Harte Maßnahmen könnten bereits vor der Prüfung beginnen (etwa Authentifizierung mit Ausweis, was ja auch bei Präsenzprüfungen rechtlich unproblematisch sei). Während der Online-Prüfung könnten eine Videoaufsicht – diese dürfte (anders als gegebenenfalls eine KI-gesteuerte Gesichtsbewegungserkennungssoftware) auch datenschutzrechtlich grundsätzlich unproblematisch sein – oder Lockdown-Browser zum Einsatz kommen.

Besonders diskutiert wurde im Anschluss an das zweite Referat die Möglichkeit, im Weg eines (stichprobenartigen) Interviews mit Prüflingen, herauszufinden, ob diese wirklich die eingereichte Leistung erbrachten haben. Schon im Impulsreferat hatte Dr. *Dieterich* hervorgehoben, dies sei nicht gänzlich unproblematisch, da jedenfalls vermieden werden müsse, dass das Interview zu einem zweiten Teil der Prüfung werde (sofern die Prüfungsordnung die Prüfung nicht ausdrücklich dergestalt zweiteilig ausgestalte). Auch stelle sich die Frage, was folge, wenn ein Prüfling sich weigere, an einem solchen Interview teilzunehmen. Es sei grundsätzlich nicht zu beanstanden, wenn die Prüfungsbehörde dies als Verdachtsmoment zum Anlass nehme, sich die eigentliche Prüfungsleistung näher anzusehen.

4. Neue Formate – modifizierte Maßstäbe?

Zum Abschluss des zweiten Teils ging Dr. *Dieterich* auf eine umfassendere Entwicklung ein, die nicht nur Online-Klausuren, sondern überhaupt innovative Prüfungsformate betreffe: Anforderungen, etwa an das Kennzeichnen von genutzten Quellen, die für traditionelle Prüfungsformate wie ‚klassische‘, mehrwöchige Hausarbeiten gelten (siehe oben I. 5.), seien nicht ohne Weiteres etwa auf *open book*-Klausuren¹¹ oder auf Power-Point-Vortragsfolien¹² übertragbar.

¹¹ Vgl. VG Dresden Beschluss vom 16. Februar 2021 – 5 L 5/21 – beck-online Rn. 47 ff.

¹² Vgl. VG Berlin, Urteil vom 30. Oktober 2024 – 12 K 195/23 – juris Rn. 32 ff.

III. Einsatz von künstlicher Intelligenz bei Prüfungen

Gleich zu Beginn des dritten Impulsvortrags hob *Fischer* hervor, beim Einsatz von KI in Prüfungen seien durchaus noch einiges im Fluss. Richter würden zudem am „Ende der Werkbank“ sitzen. Umso stärker trat die Gestaltungsperspektive der Prüfungsbehörden in den Vordergrund; *Fischer* forderte später: „Machen Sie sich Gedanken!“

1. Vielfalt des Einsatzes von KI durch Prüflinge und Prüfer

Grundsätzlich könnten bestimmte Einsatzformen von KI durch Prüflinge unter den allgemeinen Täuschungstatbestand (s.o. I. 1.) subsumiert werden. Freilich zeigte *Fischer* auf, dass es auf die konkrete Prüfungsart, Prüfungsaufgabe und die Art des Einsatzes von KI (welcher neben Textgenerierung etwa auch Textkorrektur, Paraphrasieren, Übersetzen oder Recherche betreffen könne) ankomme:

Wenn ein Prüfling beispielsweise in einer Klausur nicht hierfür zugelassene KI nutze, so setze er ein unerlaubtes Hilfsmittel ein und täusche. Insoweit sei KI freilich kein spezifisches Problem von Präsenzklausuren, da dort in der Regel ohnehin jegliche Smartphonenuutzung generell ausgeschlossen sei. Bei dem Einsatz von KI in Hausarbeiten (samt unbeaufsichtigter *open book*-,Klausuren‘) könne eine Täuschung vorliegen, wenn KI entweder (anders als Publikationen) bereits per se nicht zu den erlaubten Hilfsmitteln gehöre (siehe sogleich Ziffer 2.), oder – falls ihr Einsatz grundsätzlich erlaubt (oder gar erwünscht) sei – im Fall der mangelhaften Offenlegung (siehe unten Ziffer 3.). Bereits nach den gängigen Prüfungsordnungen dürfte jedenfalls die ungekennzeichnete Übernahme von KI-generierten Textbausteinen in Hausarbeiten eine Täuschung darstellen, wobei hinsichtlich der Feststellung, ob die Gesamtprüfungsleistung davon geprägt sei, grundsätzlich auf die allgemeine Dogmatik zu Plagiaten (siehe oben I. 5.) zurückzugreifen sei. Auch eingedenk der verfassungsrechtlichen Forderungen des Vorbehalt des Gesetzes und der Wesentlichkeitslehre seien spezielle Regelungen zur Erfassung von KI-Täuschungen grundsätzlich nicht zwingend. Es könne freilich angezeigt sein, die konkrete Prüfungsordnung auf Klarstellungspotential (etwa bei Eigenständigkeitserklärungen) zu untersuchen.

Allerdings sei Vorsicht geboten, den Täuschungstatbestand hier nicht zu überdehnen: Eine Nutzung von KI, die sich lediglich im Vorfeld des Abfassens der Prüfungsleistung, etwa als Inspirationsquelle, bewege, könne noch nicht ohne Weiteres als Täuschung erfasst werden. Auch in der analogen Welt sei das Umsehen in der Bibliothek ebenso wie der Austausch mit Kommilitonen schließlich erlaubt.

Die Dozenten gingen knapp auf den Einsatz von KI durch Prüfer ein. Dieser könne bereits bei der Erstellung der Aufgaben erfolgen, wenn KI zur Inspiration oder Kontrolle diene. Auch sei es in Grenzen zulässig, wenn Prüfer KI bei der Vorbereitung ihrer Bewertung, zum Beispiel zum Filtern von Schlagwörtern, einsetzen. Freilich müsse der Prüfer die Leistung letztlich selbst bewerten.

2. Zum Modus „Prüfling soll KI nicht nutzen“

In Betreff auf die Verhinderung unerwünschten Einsatzes von KI durch Prüflinge verwies *Fischer* im Ausgangspunkt auf die zuvor näher behandelten sanften und harten Täuschungspräventionsmaßnahmen bei Online-Prüfungen (s.o. II. 3.). Denkbar sei auch ein Ausweichen auf Aufsichtsarbeiten in Präsenz und mündliche Prüfungen oder aber die Kombination von Hausarbeiten mit einer Verteidigung. Gegebenenfalls führe es aber auch zum Ziel, die Inhalte häuslich zu bearbeitender Prüfungsaufgaben zuzuschneiden auf die – der KI gegebenenfalls unzugänglichen – besonderen Inhalte, Quellen, Techniken und Methoden der konkreten Lehrveranstaltung, wobei hier aber der Fortschritt der KI im Auge zu behalten sei.

Dr. *Dieterich* stellte die beiden bisher ergangenen Judikate zur Sanktionierung des Einsatzes von KI in prüfungsähnlichen Situationen vor.¹³ Er skizzierte die Entscheidung des Verwaltungsgerichts München, welche die Einordnung des Einsatzes von KI bei einem Essay im Rahmen eines Eignungsfeststellungsverfahrens zu einem Masterstudium als Täuschung wegen Nutzung eines unerlaubten Hilfsmittels bestätigte.¹⁴ Das Hauptaugenmerk legte er auf die zeit- und ressourcenintensive Herleitung der richterlichen Überzeugung (allgemein oben I. 2.), welche sich aus komplexen vergleichenden Betrachtungen (Vergleiche mit typischen KI-generierten Texten, mit Essays der Mitbewerber, mit wissenschaftlichen Veröffentlichungen, mit einem früheren Essay des Bewerbers) und Auswertungen schriftlicher Stellungnahmen zweier Wissenschaftler speiste. Kritik übte Dr. *Dieterich* an der Entscheidung, soweit diese den Anscheinsbeweis (allgemein oben I. 2.) bemühe. Angesichts des eher unsicheren Erfahrungswissens mit KI sei es treffender, auf die Bildung der richterlichen Überzeugung aus dem Gesamtergebnis des Verfahrens abzustellen.

Da Dr. *Dieterich* erwähnt hatte, das Gericht habe auch das Ergebnis einer Prüfsoftware in seine Würdigung einfließen lassen, fragte in der späteren Diskussion ein Teilnehmer „positiv

¹³ Vgl. VG München, Beschlüsse vom 28. November 2023 – M 3 E 23.4371 –; und vom 8. Mai 2024 – M 3 E 24.1136 – jeweils juris.

¹⁴ Vgl. VG München, Beschluss vom 28. November 2023 – M 3 E 23.4371 – juris Rn. 34, 48.

überrascht“, ob Gerichte tatsächlich auch Software zur KI-Detektion heranziehen würden. *Fischer* relativierte: Dies sei in der besprochenen Entscheidung nur ein eher nebensächlicher Punkt gewesen. Ähnlich wie bei Plagiatssoftware komme es letztlich darauf an, ob das Gericht nach eigener Prüfung überzeugt sei, dass der Prüfling eine Fehlvorstellung über die Eigenständigkeit geweckt habe. Andere Teilnehmer kamen in der Diskussion auf die Bedeutung von typischen Strukturen KI-generierter Texte sowie von archivierten früheren Prüfungsleistungen des Prüflings für die Ermittlung von verdächtigen Leistungs- und Stilsprüngen zu sprechen.

3. Zum Modus ‚Prüfling darf oder soll KI nutzen‘

Dr. *Dieterich* widmete sich sodann Möglichkeiten, die Nutzung von KI durch Prüflinge in Prüfungen einzubeziehen. Auf didaktischer Ebene könne das durchaus fruchtbar sein. So sei etwa an KI-bezogene Aufgaben wie die Formulierung von Prompts oder die kritische Würdigung von KI-generierten Ergebnissen zu denken. Vorstellbar sei, dass Prüfer die Qualität des Umgangs mit KI (gute *prompts*, Effizienz, Kontrolle, Reflexion) als Bewertungskriterium heranziehen. Auch bei herkömmlichen Prüfungsformaten, etwa bei Hausarbeiten, könne die Nutzung von KI zugelassen werden, was gegebenenfalls eine Änderung der Prüfungsordnung erforderlich machen könnte. In der späteren Diskussion stellten die Dozenten klar, dass unter Umständen auch eine Änderung der Hilfsmittelbekanntmachung genügen könne oder solche Entscheidungen den einzelnen Prüfern überantwortet werden könnten.

Sei KI-Nutzung zulässig, so komme es grundsätzlich auf die hinreichend gekennzeichnete Nutzung der KI durch die Prüflinge an. Es sei aber auch vorstellbar, dass in der konkreten Aufgabenstellung – gewissermaßen im Geist von ‚enhanced games‘ – sogar auf eine solche Kennzeichnung verzichtet werde. Bereits zuvor hatte *Fischer* Ausführungen zu den Möglichkeiten und Grenzen der Kennzeichnung von KI-Nutzung gemacht. Grundsätzlich seien KI-generierte Inhalte keine zitierfähige Quelle im klassischen Sinn, unter anderem mangels konkret identifizierbarer Urheber und Stabilität der Antworten. Dennoch sei eine Kennzeichnung der IT-Nutzung vorstellbar, zum Beispiel durch die Dokumentation der KI-Hervorbringungen in Anhängen zur erstellten Arbeit. Gegebenenfalls biete es sich an, wenn Hochschulen ihre Satzungen zur guten wissenschaftlichen Praxis ergänzten. Zur Bestimmung der Dichte der Kennzeichnungspflicht biete sich grundsätzlich ein Rückgriff auf die Formel, wonach „der Leser an jeder Stelle [wissen können müsse], wer zu ihm spricht“ (s. oben I. 5.). an. Spannend war die Frage einer Teilnehmerin, ob – im Fall der grundsätzlichen Erlaubnis von KI-Verwendung – eine Täuschung des Prüflings anzunehmen sei, wenn dieser eine von der KI

halluzinierte Quelle übernehme – oder ob dies schlicht auf der Bewertungsebene zu berücksichtigen sei. Nach den Dozenten dürfte dies von der Aufgabenstellung sowie davon abhängen, ob der Prüfling (konkludent) erklärt habe, die angegebenen Quellen selbstständig ausgewertet und insoweit wissenschaftlich gearbeitet zu haben.

Die Diskussion kam schließlich auch auf das Problem zu sprechen, dass einzelne Prüflinge wegen ihrer überdurchschnittlichen finanziellen Möglichkeiten in der Lage sein könnten, sich Zugang zu besonders leistungsstarken, aber teuren KI-Lösungen zu verschaffen. Dr. *Dieterich* gab zunächst zu bedenken, dass das Prüfungsrecht als solches grundsätzlich blind für sozio-ökonomische Unterschiede sei, und wies auf die Verbreitung kommerzieller Repetitorien in manchen Disziplinen hin. Die Dozenten waren sich jedoch einig, dass dann, wenn nicht nur die Vorbereitung auf Prüfungen, sondern die Möglichkeiten bei der Leistungserbringung selbst in nennenswertem Maß betroffen seien, die Chancengleichheit gefährdet werde. Prof. *Ulf Pallme König* und Prof. Dr. *Volker Epping* hoben sodann zum Abschluss der Tagung auch nochmals die Gestaltungs- und Versorgungsverantwortung der Prüfungsbehörden, vor allem der Hochschulen, hervor. Dazu gehöre neben einer umfassenden Fortbildung des Personals auch – natürlich im Rahmen des Machbaren – das Vorhalten einer KI-Grundversorgung.

Dr. *Rico David Neugärtner*, LL.M. (Cornell) ist Richter und als Beisitzer in einer für Prüfungsrecht zuständigen Kammer des Verwaltungsgerichts Berlin tätig.