

# Sebastian Dötterl

## *Der Elefant im Raum – Generative Künstliche Intelligenz und die Zukunft der juristischen Ausbildung<sup>1</sup>*

### Übersicht

#### I. Ausgangssituation: generative KI verändert den Rechtsmarkt

1. Können große Sprachmodelle Jura?
2. Veränderung der juristischen Arbeitswelt
3. KI wird nicht nur „Routineaufgaben“ übernehmen
  - a) Der Weg zur Expertise führt über die Entwicklung von Routinen
  - b) KI beschränkt sich nicht auf Routineaufgaben
  - c) Konsequenzen für das „Training on the job“
4. Das skillset der Juristen der Zukunft

#### II. Lernen „über KI“: Konsequenzen für den Stoffplan

1. Mehr Interdisziplinarität
2. Mehr Recht der Digitalisierung

#### III. Lernen und Lehren „mit KI“

1. Ein Werkstattbericht
2. Neue Möglichkeiten in der Lehre
  - a) KI könnte klassische Lehrformate bereichern
  - b) Weitergehend kommen auch neue Formate in Frage
3. Neue Möglichkeiten für das Lernen
4. Probleme und Fallstricke
  - a) Mögliche Einschränkung der Fähigkeit zu eigenständigem, kritischem Denken
  - b) Entmenschlichung der Ausbildung
  - c) Rechtliche Aspekte
  - d) Moralische Aspekte
  - e) Zukünftige Kosten und Monopole
5. Zwischenfazit

#### IV. Prüfen „durch KI“?

1. Digitalisierung als Grundvoraussetzung
2. Welche Rolle kann KI in Prüfungen spielen?
  - a) Klassische Klausuren
  - b) Haus- und Seminararbeiten

c) Prüfungen, bei denen KI-Einsatz erlaubt ist

d) Mündliche Prüfungen

e) Multiple-Choice

#### 3. Bewertung durch KI

#### V. Zusammenfassung

Die juristische Berufswelt befindet sich in einem tiefgreifenden Wandel. Sinkende Eingangszahlen bei den Zivilgerichten<sup>2</sup>, immer weniger klassische Einzelkanzleien<sup>3</sup>, die zunehmende Komplexität, Internationalisierung und Spezialisierung des Rechts sowie die Digitalisierung haben bereits vor Jahren einen Veränderungsprozess eingeleitet. Mit dem Bekanntwerden generativer Künstlicher Intelligenz (KI) in der breiten Öffentlichkeit seit Ende 2022 hat diese Entwicklung eine neue Dimension erreicht.

In nur wenigen Jahren haben diese Systeme einen Leistungsstand erreicht, der sie bestenfalls zu hilfreichen Werkzeugen, schlimmstenfalls zu Jobkillern für menschliche Juristen macht. Angesichts dieser oft als *Gamechanger*<sup>4</sup> bezeichneten Entwicklung steht die juristische Ausbildung vor der Herausforderung, die nächste Juristengeneration auf ihre berufliche Zukunft vorzubereiten.

Zentral sind dabei folgende Fragen:

- Welche Veränderungen ergeben sich durch generative KI für die juristischen Berufsbilder?
- Welches Wissen und welche Kompetenzen benötigen angehende Juristen zukünftig?
- Wie können KI-Tools sinnvoll in Jurastudium und Referendariat integriert werden?
- Welche Auswirkungen hat generative KI auf juristische Prüfungsformate?

1 Der Beitrag gibt die persönliche Meinung des Verfassers wieder. Besonderer Dank gebührt Christine Freifrau von Massenbach und Dr. Ann-Kristin Mayrhofer für wertvolle Hinweise. Verbleibende Irrtümer sind ausschließlich solche des Verfassers.

2 Vgl. Meller-Hannich et al., Abschlussbericht zum Forschungsvorhaben „Erforschung der Ursachen des Rückgangs der Eingangszahlen bei den Zivilgerichten“ (2023), [bericht\\_Eingangszahlen\\_Zivilgerichte.html. Alle Internetquellen wurden letztmals am 14.5.2025 abgerufen.](https://www.bmj.de/SharedDocs/Downloads/DE/Fachinformationen/Abschluss-</a></p></div><div data-bbox=)

3 Vgl. Meldung „Neue Zahlen zum Anwaltsmarkt: Der Abwärtstrend setzt sich fort“ vom 28.5.2024, <https://rsw.beck.de/aktuell/daily/meldung/detail/anwaltsmarkt-weniger-rechtsanwaeltmehrunternehmensjuristen-anwaeltinnen-frauen>.

4 Vgl. etwa das Motto des 33. EDV-Gerichtstags 2024: „Recht im Umbruch – KI als Gamechanger?“.

Der vorliegende Beitrag erhebt keinen Anspruch, diese Fragen vollständig zu beantworten, sondern versucht, in der Reformdiskussion den Blick auf generative KI als den bislang nicht ausreichend thematisierten „Elefanten im Raum“ zu lenken. Nach einer Analyse der Ausgangssituation (I.) werden Konsequenzen für den juristischen Stoffplan (II.) überlegt. Anschließend werden die Potenziale und Risiken von KI-Tools im Studium diskutiert (III.), bevor die Auswirkungen auf juristische Prüfungsformate untersucht werden (IV.).

## I. Ausgangssituation: Generative KI verändert den Rechtsmarkt

### 1. Können Große Sprachmodelle Jura?

Große Sprachmodelle (Large Language Models, LLMs) beeindrucken durch hohe Textqualität und gutes Textverständnis. Ihre Leistungsfähigkeit im juristischen Kontext wurde bereits in verschiedenen Studien untersucht.<sup>5</sup> Das mittlerweile veraltete Modell GPT-4 würde das anspruchsvolle US-amerikanische Bar Exam bestehen.<sup>6</sup> LLMs agierten bei der Prüfung umfangreicher Verträge qualitativ auf Augenhöhe mit Berufseinsteigern, und das in einem Bruchteil der Zeit und zu einem Bruchteil der Kosten.<sup>7</sup> Studierende, die in Prüfungsaufgaben ein Reasoning-Modell<sup>8</sup> bzw. ein spezielles RAG-Modell<sup>9</sup> nutzten, waren schneller und schnitten besser ab als ohne KI-Unterstützung.<sup>10</sup> Damit erfolgte ein Quantensprung gegenüber einer Vorgängerstudie mit dem alten Modell GPT-4, in der die Studierenden nur schneller, jedoch nicht wesentlich besser wurden.<sup>11</sup>

In einer weiteren Untersuchungsreihe wurden identische Fallmaterialien aus einem Kriegsverbrecherprozess

sowohl erfahrenen Richtern als auch einem KI-System zur Entscheidung vorgelegt. Es zeigte sich, dass die KI sich stärker an den vorgegebenen Präzedenzfällen orientierte. Die menschlichen Richter nahmen sich, abhängig von der Charakterisierung des Angeklagten als sympathisch oder unsympathisch, mehr Freiheiten bei ihrer Entscheidung.<sup>12</sup>

Auf der anderen Seite sind Halluzinationen – also falsche Antworten – weiterhin ein Problem.<sup>13</sup> Schlagzeilen machte ein Fall, in dem ein Rechtsanwalt von ChatGPT erfundene Präzedenzfälle bei einem New Yorker Gericht einreichte.<sup>14</sup> Dieser exemplarische Fall lenkt den Blick auf das allgemeine Problem, dass Fragen zu juristischem Faktenwissen von großen Sprachmodellen regelmäßig falsch beantwortet werden.<sup>15</sup> Auch die Verwendung von RAG kann Halluzinationen bei der rechtlichen Recherche nicht vollständig verhindern.<sup>16</sup>

Alle genannten Untersuchungen basieren auf englischsprachigen Aufgabenstellungen und *case law* Rechtssystemen. Die Ergebnisse lassen sich daher nicht ohne Weiteres auf die deutsche Sprache und das deutsche Recht übertragen, da weniger deutschsprachiges Trainingsmaterial zur Verfügung steht. Ein Preprint hat gezeigt, dass LLMs bei deutschsprachigen juristischen Aufgabenstellungen signifikant schlechter abschneiden als bei englischsprachigen Aufgaben.<sup>17</sup>

Zudem sind dies nur Momentaufnahmen, da laufend neue Sprachmodelle veröffentlicht werden und zunehmend das Training bzw. die Verknüpfung mit spezifischem juristischem Quellenmaterial hinzutritt. Aufgrund der schnellen technologischen Entwicklung ist zu erwarten, dass die Leistungsfähigkeit der KI-Systeme sich weiter steigern lässt, etwa durch die relative neue

5 Siehe hierzu Dötterl, *LTZ* 2024, 235.

6 Katz et al, GPT-4 passes the bar exam (2024); *Phil. Trans. R. Soc. A* 382: 20230254, <https://doi.org/10.1098/rsta.2023.0254>. Umstritten ist, wie gut GPT-4 tatsächlich abschneidet, vgl. Martinez, Reevaluating GPT-4's bar exam performance, *Artif. Intell. Law* (2024), <https://doi.org/10.1007/s10506-024-09396-9>.

7 Martin et al., Better Call GPT, Comparing Large Language Models Against Lawyers (2023), <https://arxiv.org/html/2401.16212v1>, zur Methodik siehe Dötterl (Fn. 5), 237.

8 Reasoning-Modelle sind Sprachmodelle, die darauf trainiert wurden, komplexe Probleme durch eine schrittweise Zerlegung zu lösen. Diese Modelle geben typischerweise ihre vermeintlichen Zwischenschritte im Denk- und Analysevorgang aus (sog. *Chain-of-Thought*).

9 RAG steht für *Retrieval Augmented Generation*; damit ist im vorliegenden Kontext der Zugriff der Sprachmodelle auf kuratierte juristische Daten gemeint.

10 Schwarcz et al., AI-Powered Lawyering: AI Reasoning Models, Retrieval Augmented Generation, and the Future of Legal Practice, *Minnesota Legal Studies Research Paper No. 25-16* (2025), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5162111](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5162111).

11 Choi/Monahan/Schwarcz, *Lawyering in the Age of Artificial Intelligence*, 109 *Minn. L. Rev.* 147, 150 (2024), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4626276](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4626276).

12 Posner/Saran, Judge AI: Assessing Large Language Models in Judicial Decision-Making (2025), <http://dx.doi.org/10.2139/ssrn.5098708>.

13 Jaźwińska/Chandrasekar, AI Search Has A Citation Problem (2025), [https://www.cjr.org/tow\\_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php](https://www.cjr.org/tow_center/we-compared-eight-ai-search-engines-theyre-all-bad-at-citing-news.php).

14 LTO-Meldung vom 30.5.2023: „Anwalt fällt auf Fake-Urteile von ChatGPT rein“; <https://www.lto.de/recht/kurioses/k/anwalt-new-york-chatgpt-recherche-schriftsatz-fake-urteile>.

15 Dahl et al., Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models, *Journal of Legal Analysis*, Volume 16, Issue 1, 2024, S. 64–93, <https://doi.org/10.1093/jla/laae003>.

16 Magesh et al., Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools (2024), <https://arxiv.org/abs/2405.20362>.

17 Fan et al., LEXAM: Benchmarking Legal Reasoning on 340 Law Exams (2025), S. 7, 25, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5265144](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5265144).

"Deep Research"-Funktion einiger Modelle, oder indem die aus Wahrscheinlichkeit beruhende Texterzeugung mit regelbasierten Ansätzen (sog. hybride KI)<sup>18</sup> oder mit anderen Arten von KI-Algorithmen kombiniert wird (compound AI).<sup>19</sup> Andererseits ist nicht zu erwarten, dass das Problem der Halluzinationen verschwindet: Das aktuelle Reasoning-Modell o3 von Open.AI halluziniert sogar öfter als die Vorgängermodelle.<sup>20</sup>

## 2. Veränderung der juristischen Arbeitswelt

Der Geist ist aus der Flasche: Generative KI wird bleiben, insbesondere im juristischen Bereich. Es sind zahlreiche auf den Rechtsmarkt zugeschnittene Produkte auf dem Markt, die auf diese Technologie setzen. Dies betrifft etwa speziell mit juristischen Texten trainierte Sprachmodelle,<sup>21</sup> KI-gestützte Rechtsrecherche,<sup>22</sup> oder Kanzleisoftware, die Schriftsatzentwürfe anhand der Aktenlage erstellen kann.<sup>23</sup> Tools erstellen bereits angepasste Vertragsentwürfe oder überprüfen Vertragsklauseln auf ihre Wirksamkeit.<sup>24</sup> Somit hält generative KI Einzug in immer mehr Anwaltskanzleien und Rechtsabteilungen.

Diese Entwicklung wird sich voraussichtlich fortsetzen und früher oder später auch den staatlichen Bereich inklusive Rechtsprechung erreichen. In der deutschen Justiz werden derzeit erste Erfahrungen mit generativer KI gesammelt.<sup>25</sup> Im chinesischen Shenzhen ist eine spe-

zialisierte generative KI bereits im Praxiseinsatz: Sie identifiziert anhand der Verfahrensakten streitige Punkte, schlägt Fragen für die mündliche Verhandlung vor und entwirft die schriftliche Urteilsbegründung auf Grundlage der zuvor von der Richterin getroffenen Entscheidung.<sup>26</sup>

Auch gesellschaftlich verschiebt sich etwas. In einer britischen Studie bevorzugten Laien – sofern sie die Herkunft nicht kannten – den Rechtsrat einer KI gegenüber demjenigen einer Anwältin.<sup>27</sup> Laut einer Umfrage in Deutschland würden 15 Prozent der Befragten bei rechtlichen Problemen auf die Hilfe einer Anwältin verzichten und stattdessen etwa eine KI nutzen.<sup>28</sup>

Ob diese Entwicklung irgendwann auch dazu führen könnte, dass Menschen Urteile lieber von einer KI erhalten wollen als von einer menschlichen Richterin, ist nicht absehbar. Ein Startup reklamiert jedenfalls schon jetzt, eine KI-basierte Schiedsstelle entwickelt zu haben, die effizienter, kostengünstiger und zuverlässiger als menschliche Schiedsrichter arbeite und perspektivisch staatliche Gerichte funktional überflüssig machen könne.<sup>29</sup>

## 3. KI wird nicht nur „Routineaufgaben“ übernehmen

Die zunehmende Integration von KI-Systemen in den Rechtsbereich wird häufig mit dem Narrativ verbunden, dass künstliche Intelligenz nur Routineaufgaben über-

18 Vgl. *Dirschl/Hartz* in Hinkelmann et al (Hrsg.), *Hybride KI mit Machine Learning und Knowledge Graphs* (2025), S. 349 ff.; [https://link.springer.com/chapter/10.1007/978-3-658-44781-6\\_15](https://link.springer.com/chapter/10.1007/978-3-658-44781-6_15); auf diesem Ansatz basiert auch die Software der Rulemapping Group, die im April 2025 eine Finanzierungsrunde über 12 Millionen Euro, u.a. mit der Bundesagentur für Sprunginnovationen (SPRIND) abgeschlossen hat, vgl. <https://rulemapping.com/de/stories/funding>.

19 *Eder*, KIR 2025, 1 (2); abrufbar unter <https://rsw.beck.de/zeitschriften/kir/editorial/2025/02/13/transformation-der-juristischen-ausbildung-im-zeitalter-von-ki--innovation-durch-hackathons>.

20 Vgl. OpenAI o3 und o4-mini System Card, S. 4, <https://cdn.openai.com/pdf/2221c875-02dc-4789-800b-e7758f3722c1/o3-and-o4-mini-system-card.pdf>.

21 Ein Beispiel in Deutschland ist das Modell Noxtua, in dessen Weiterentwicklung nun der Verlag C.H. Beck maßgeblich investiert, siehe Pressemeldung des Verlags C.H. Beck vom 23.4.2025, <https://presse.beck.de/pressemitteilungen-archiv/2025/chbeck-wird-fuehrender-investor-bei-rechts-ki-noxtua/>; international bekannt ist das Produkt Harvey (<https://www.harvey.ai>).

22 Ein Beispiel aus Deutschland ist „Frag den Grüneberg“, siehe Pressemeldung des Verlags C.H. Beck vom 5.12.2024, <https://presse.beck.de/pressemitteilungen-archiv/2024/frag-den-grueneberg-die-funktionen-des-chat-book-im-uberblick>.

23 Beispiele aus Deutschland sind etwa Libra (<https://libratech.ai/de>) oder JUNE (<https://www.june.de>).

24 Ein international erfolgreiches Beispiel ist Henchman (<https://henchman.io>).

25 Vgl. zum Generativen Sprachmodell für die Justiz (GSJ) die Pressemitteilung des Bayerischen Staatsministeriums der Justiz vom 31.07.2024: <https://www.bayern.de/kuenstliche-intelligenz-im-einsatz-fuer-die-justiz-start-fuer-gemeinsames-forschungsprojekt-von-bayern-und-nrw-generatives-sprachmodell-soll-richterinnen-und-richter-entlasten-wissenschaftliche/>; zur Massenverfahrens-Assistenz (MAKI), der Erkenntnismittelassistentz EMIL und der Tatbestandsassistentz für Asylgerichtsverfahren TabEA siehe <https://anwaltspraxis-magazin.de/fachbeitraege/elektronischer-rechtsverkehr/2025/02/09/b-ki-in-der-niedersaechsischen-justiz-80/>.

26 *Liu/Li*, How do judges use large language models? Evidence from Shenzhen (2024), *Journal of Legal Analysis*, Volume 16, Issue 1 2024, S. 235 ff, <https://doi.org/10.1093/jla/laee009>.

27 *Schneiders* et al., Objection Overruled! Lay People can Distinguish Large Language Models from Lawyers, but still Favour Advice from an LLM (2025), <https://arxiv.org/abs/2409.07871>.

28 Siehe Presseinformation der Bitkom vom 09.12.2024: „Online-Tool statt Rechtsanwalt: Jeder Siebte würde sich lieber digital helfen lassen“, <https://www.bitkom.org/Presse/Presseinformation/Online-Tool-statt-Rechtsanwalt>.

29 *Kieffaber/Gandall/McLaren*, We Built Judge. AI. And You Should Buy It (2025), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5115184](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5115184).

nehmen wird, während sich menschliche Juristen auf komplexere und strategische Aufgaben konzentrieren können.<sup>30</sup> Diese scheinbar plausible Arbeitsteilung suggeriert eine Win-Win-Situation: Befreiung von monotonen Tätigkeiten und mehr Zeit für Wesentliches einerseits, steigende Effizienz und sinkende Kosten andererseits. Bei näherer Betrachtung entpuppt sich diese beschwichtigende Sichtweise jedoch als falsch.

#### a) Der Weg zur Expertise führt über die Entwicklung von Routinen

Der erste Denkfehler liegt in der Unterschätzung der Bedeutung von Routineaufgaben für den Kompetenzerwerb. Bevor für Nachwuchsjuristen etwas „Routine“ wird, muss es zunächst mühsam erlernt werden. Recherchieren oder das Erstellen von Entwürfen mögen als Pflichtübungen erscheinen, vermitteln aber zugleich auch Fähigkeiten wie Detailgenauigkeit, Erkennen rechtlicher Probleme und Absichten einer Fülle an Informationen.

Wenn KI-Systeme diese Aufgaben übernehmen, wird der Mensch nicht nur von einer – vielleicht unliebsamen – Aufgabe befreit, sondern verliert auch die damit verbundene Lernerfahrung. Ohne den vorherigen Erwerb von Routine werden Nachwuchsjuristen nicht in der Lage sein, auf Anhieb „komplexere und strategische“ Rollen auszufüllen.

#### b) KI beschränkt sich nicht auf Routineaufgaben

Der zweite Irrtum liegt in der Annahme, dass KI lediglich Routineaufgaben übernehmen kann und wird. Die technologische Entwicklung zeigt in eine andere Richtung: KI-Systeme – denen juristischer Berufsstolz und -ethos herzlich egal ist – klettern die Wertschöpfungskette hinauf und dringen schon jetzt in den Kernbereich juristischer Tätigkeit vor, etwa die Texterstellung, Vertragsgestaltung<sup>31</sup> oder die Auslegung und Subsumtion von Gesetzen.<sup>32</sup>

#### c) Konsequenzen für das „Training on the job“

Diese Entwicklung hat Auswirkungen auf den Berufseinstieg. Wenn KI klassische Tätigkeiten von Berufsein-

steigern zu einem Bruchteil der Kosten übernehmen kann, werden Mandanten nicht bereit sein, hohe Stundensätze zu bezahlen, damit eine junge Juristin diese Aufgaben erledigt. Dies dürfte zu einer grundlegenden Veränderung der Ausbildungs- und Beschäftigungsstrukturen in juristischen Berufen führen.<sup>33</sup> Empirisch zu belegen ist dies derzeit allerdings nicht.<sup>34</sup>

#### 4. Das Skillset der Juristen der Zukunft

Diese Entwicklungen führen zu einem paradoxen Bildungsauftrag: Angehende Juristen müssen Zeit, Energie und intellektuelle Anstrengung in das Erlernen von Fähigkeiten investieren, die KI-Systeme schon jetzt – und erst recht in Zukunft – besser beherrschen. Doch nur durch diesen Prozess werden sie tiefergehendes Verständnis und die nötige Urteilsfähigkeit entwickeln, um KI-Systeme später „bändigen“ und deren Fehler erkennen und korrigieren zu können.

Große Sprachmodelle können die juristische Methode zwar täuschend echt und teilweise auch in beeindruckender Qualität imitieren. Dies kann aber nicht darüber hinwegtäuschen, dass es sich dabei nur um Mimikry der juristischen Methodik handelt. Die Systeme haben kein Judiz, kein Verständnis von der Welt und können Werte und Interessen nicht abwägen.<sup>35</sup> KI-Systeme erstellen anhand von Wahrscheinlichkeitsrechnungen plausibel klingende Texte und hängen dabei von Trainingsdaten aus der Vergangenheit ab. Somit dürfte die KI überfordert sein, wenn ein Fall nicht durchs Trainingsmaterial abgedeckt ist, weil z.B. ein unwahrscheinliches Ereignis zu beurteilen ist.<sup>36</sup> Zum anderen könnte vorhandenes Trainingsmaterial auch veraltet (z.B. aufgehobene Gesetze, überholte Entscheidungen) oder schlicht unrichtig sein.

Damit ist klar: Gute Juristen werden auch in Zukunft nicht überflüssig. Ganz im Gegenteil: Sie werden dringend gebraucht. Ihnen wird es – in welcher Rolle auch immer – vorbehalten sein, rechtliche Entscheidungen zu treffen, und für *ihre* Entscheidungen auch die Verantwortung zu übernehmen.

Nötig dafür bleiben die klassischen juristischen Fähigkeiten. Analytisches Denken, Methodik, Problembe-

30 So z.B. *Ajevski et al.*, ChatGPT and the future of legal education and practice. *The Law Teacher*, 57(3), 352 (2023), bei Fn. 26; <https://doi.org/10.1080/03069400.2023.2207426>.

31 Vgl. oben Fn. 23 f.

32 *Dietrich*, NJW 2024, 2092.

33 *Eder* (Fn. 19), S. 1.

34 *Humlum/Vestergaard*, Large Language Models, Small Labor Market Effects (2025), S. 30, 49; <https://ssrn.com/abstract=5219933>.

35 Ferner fehlen KI-Systemen genuin menschliche Eigenschaften wie Empathie, Verhandlungsgeschick, Organisationstalent und vieles mehr.

36 *Meder*, JZ 2023, 1041 (1042, 1047).

wusstsein und Falllösungskompetenz. Die Stärke des traditionellen Studiums bei der Vermittlung der juristischen Methode muss daher unbedingt bewahrt werden.

Diese große Stärke allein reicht in der modernen Welt aber nicht aus. Juristen benötigen in Zukunft ein noch breiteres und anspruchsvolleres Skillset.<sup>37</sup>

## II. Lernen „über KI“: Konsequenzen für den Stoffplan

Juristen benötigen künftig zusätzliche Kompetenzen. Dies erfordert eine kritische Überprüfung und Erweiterung der Lehrinhalte sowohl im Studium als auch im Referendariat in zweierlei Hinsicht: Zum einen sollten interdisziplinäre Bezüge des Lehrplans vor allem in technischer Sicht ausgebaut werden, zum anderen bedarf es einer verstärkten Integration digitalrechtlicher Fragestellungen.

Erfreulicherweise haben bereits viele deutsche Universitäten ihr Lehrangebot erweitert.<sup>38</sup> In Bayern wurde 2023 im Referendariat ein neues Berufsfeld „Informatonstechnologierecht und Legal Tech“ eingeführt. Diese Initiativen sollten ausgebaut und noch stärker in den Pflichtstoff integriert werden.<sup>39</sup>

### 1. Mehr Interdisziplinarität

Juristen benötigen ein technisches Grundverständnis für KI: Funktionsweise, Stärken, Schwächen. Gefordert sind Basiskenntnisse über Coding und Statistik, ohne dass Juristen gleich zu Informatikern ausgebildet werden müssen.<sup>40</sup>

Auch Kenntnisse über typische menschliche Fehler in juristischen Denk- und Entscheidungsprozessen<sup>41</sup> sind nicht nur für die Selbstreflexion und das eigene Selbstverständnis, sondern auch für den richtigen Umgang mit KI-Systemen von Relevanz.

Zudem sollten Grundlagen des *Legal Design* in die Ausbildung integriert werden. Hier geht es darum, komplexe rechtliche Inhalte verständlich zu kommunizieren, z.B. durch Visualisierungen, Strukturierung und Alltagssprache. Generative KI hat hier Stärken, doch auch bei

der „Übersetzung“ juristischer Fachsprache braucht es Menschen, die für die Richtigkeit der Erklärungen einstehen.

### 2. Mehr Recht der Digitalisierung

Neben der interdisziplinären Erweiterung sollte der Lehrstoff inhaltlich stärker auf die Digitalisierung ausgerichtet werden. Dies betrifft einerseits digitalisierungsbezogene Fragestellungen in klassischen Rechtsgebieten. Hier besteht kein direkter Handlungsbedarf. Denn neue Fragen und Phänomene werden in Wissenschaft und Praxis virulent und damit auch in der Ausbildung behandelt. Ein Beispiel wären etwa die Änderungen des BGB infolge der Digitale-Inhalte-Richtlinie und der Warenkauf-Richtlinie. Digitale Produkte finden dadurch Eingang in den klassischen Prüfungsstoff. So wird es auch mit den Rechtsproblemen rund um KI sein.

Zunehmend wichtiger wird daneben das „Recht der Digitalisierung“, etwa Datenschutzrecht oder die Regulierung von Künstlicher Intelligenz. Hier ist die Digitalgesetzgebung der Europäischen Union von enormer praktischer Bedeutung (z.B. die Datenschutzgrundverordnung und die KI-Verordnung). Die juristische Ausbildung sollte daher noch stärker auf fundierte Kenntnis des europäischen Rechtsrahmens, besonders wichtiger Rechtsakte und der Wechselwirkungen mit nationalem Recht abzielen.

## III. Lernen und Lehren „mit KI“

Unabhängig von der Frage des Lehrstoffs bietet generative KI Potentiale, aber auch Risiken im Rahmen der juristischen Ausbildung.

### 1. Ein Werkstattbericht

Noch ist der Einfluss generativer KI auf das juristische Studium überschaubar. Gemeinsam mit *Dr. Ann-Kristin Mayrhofer* leitete der *Verfasser* 2024 und erneut 2025 ein juristisches Grundlagenseminar an der Ludwig-Maximilians-Universität München. Das Besondere an dem Seminar ist, dass der Einsatz beliebiger generativer KI-

37 Vgl. bereits *Dötterl* (Fn. 5), 239.

38 Zu nennen sind ohne Anspruch auf Vollständigkeit: das Center for Legal Technology and Data Science der Bucerius Law School, Zertifikatsstudiengänge an der Ludwig-Maximilians-Universität München und der Universität Göttingen, spezialisierte Masterstudiengänge etwa an der Universität Regensburg, der Universität Passau, der Universität zu Köln, der Universität des Saarlandes und der Universität Würzburg.

39 So auch *Ogorek/Barthle*, *LTZ* 20023, 114 (117).

40 So auch *Hartung*, *Continuing legal education in Germany – Digitalization* (2023), S. 23; [https://ink.library.smu.edu.sg/sol\\_research/4533](https://ink.library.smu.edu.sg/sol_research/4533).

41 Vgl. etwa *Guthrie et al.*, „Blinking on the Bench: How Judges Decide Cases“, 93 *Cornell Law Review* (2007); <https://scholarship.law.cornell.edu/facpub/917/>, *Risse NJW* 2018, 2848.

Tools bei der Erstellung der Arbeit nicht nur erlaubt, sondern verpflichtend ist. In einem Anhang zur Seminararbeit und in einem Abschnitt des mündlichen Vortrags erläutern die Studierenden, welche KI-Tools sie auf welche Weise genutzt und welche Erfahrungen sie dabei gemacht haben.<sup>42</sup>

Hier zeigte sich zwar 2025 gegenüber 2024, dass die frei verfügbaren KI-Tools leistungsfähiger und für mehr Aufgaben eingesetzt wurden. Gelobt wurde insbesondere, dass KI-Tools einen ersten Überblick über das Thema geben konnten, bei der Ideensammlung und der Strukturierung der Seminararbeit unterstützten und auch einzelne Abschnitte der Arbeit formulieren konnten. In der Diskussion zeigten sich die Studierenden aber auch 2025 wie schon im Vorjahr zurückhaltend. Kritisiert wurden insbesondere fehlerhafte Ausgaben der KI bei Rechtsfragen. Bei der Suche nach Literatur waren halluzinierte Quellen die Regel. Soweit die Tools existierende Quellen auffanden, waren diese oft irrelevant oder wiesen nicht die für die wissenschaftliche Arbeit nötige Qualität auf. Moniert wurde ferner die Unfähigkeit der KI-Tools zu korrektem Zitieren. Manch einer hatte das Gefühl, dass die Interaktion mit KI-Tools mehr Zeit in Anspruch nahm, als es eine klassische Recherche getan hätte.

Positiv wurde dagegen von vielen vermerkt, dass KI motiviert und psychologisch unterstützt, weil man sich nicht mit dem Thema alleingelassen fühlt und nicht bei Null anfangen muss.

Diese eher ernüchternden Erfahrungen stellen eine Momentaufnahme dar und beziehen sich im Wesentlichen auf die kostenlos verfügbaren KI-Tools. Die technische Entwicklung geht jedoch weiter.

## 2. Neue Möglichkeiten in der Lehre

Vielfach wird generativer KI das Potential zugeschrieben, juristisches Lehren und Lernen zukünftig effektiver und motivierender zu machen. Für die Lehre werden beispielsweise folgende Möglichkeiten diskutiert.

### a) KI könnte klassische Lehrformate bereichern

Zum Beispiel könnten eine in der Vorlesung oder Referendar-AG besprochene Frage live in eine KI eingegeben

und deren Antwort auf Stärken und Schwächen geprüft werden. Weitergehend könnte auch ein kompletter Fall zunächst klassisch besprochen, dann durch die KI bearbeitet und anschließend das KI-Ergebnis gemeinsam analysiert werden.<sup>43</sup>

In Workshops könnten Studierende und Referendare in kleineren Gruppen Fälle und Probleme mit Hilfe von KI-Tools bearbeiten<sup>44</sup> – eine Idee, die auch in den traditionellen begleitenden AGs und Übungen aufgegriffen werden könnte. Weitere Möglichkeiten wären etwa die Bewertung von KI-generierten Lösungen durch Studierende, und die Weiterbearbeitung und Optimierung solcher Entwürfe.<sup>45</sup>

Seminare, die auf den bewussten und kritischen Einsatz von KI-Tools abzielen, sind ebenfalls eine Möglichkeit (siehe bereits oben unter 1.).

Lehrende und Prüfende könnten zudem von generativer KI profitieren, indem sie sich Anregungen für die Vermittlung oder Prüfung bestimmter Materien geben lassen. KI könnte etwa basierend auf Vorlesungsmaterialien Multiple-Choice-Aufgaben oder Sachverhalte mit Lösungsskizze erstellen.<sup>46</sup>

### b) Weitergehend kommen auch neue Formate in Frage

KI könnte es Studierenden und Referendaren ermöglichen, praktische juristische Fähigkeiten in einer kontrollierten, aber realitätsnahen Umgebung zu entwickeln. KI-Systeme könnten als virtuelle Verhandlungspartner oder gegnerische Anwälte fungieren, die auf die Argumentation der Studierenden dynamisch reagieren. Diese Idee wurde bereits in die Praxis umgesetzt: In einem Projekt der TH Köln können Studierende bei simulierten Gerichtsverhandlungen mit KI-generierten Zeugen interagieren, die auf verschiedene Befragungstechniken unterschiedlich reagieren, wobei VR-Brillen zum Einsatz kommen können.<sup>47</sup>

In Frage kommen auch die Entwicklung von Legal Tech Tools und ein „Pitch“ der Idee nach dem Vorbild von Hackathons.<sup>48</sup> Ergänzend könnten vermehrt Moot Courts oder Case-Studies hinzukommen.<sup>49</sup> Interessant erscheint zudem die Idee eines „Flipped-Classroom“, bei dem die Studierenden, ggf. mit KI-Unterstützung, die

42 Vgl. zum Konzept *Mayrhofer*, LTZ 2024, 189; sowie Blogbeitrag der Studierendeninitiative MLTech vom 30.4.2025, <https://www.ml-tech.org/blog/einfach-mal-probieren-wie-das-grundlagenseminar-kuenstliche-intelligenz-und-recht-das-bewusstsein-fuer-ki-kompetenzen-der-studierenden-staerken-soll/>.

43 *Harnos*, im April 2025 auf LinkedIn geteiltes Lehrkonzept, [https://www.linkedin.com/posts/rafael-harnos-8b530a190\\_ki-activity-7323033115680034818-aFey](https://www.linkedin.com/posts/rafael-harnos-8b530a190_ki-activity-7323033115680034818-aFey).

44 *Harnos* (Fn. 43).

45 *Bliss*, Teaching Law in the Age of Generative AI, 64 JURIMETRICS J. (2024), S. 111 (153 ff), <https://www.americanbar.org/>

[content/dam/aba/publications/Jurimetrics/winter-2024/teaching-law-in-the-age-of-generative-ai.pdf](https://www.americanbar.org/content/dam/aba/publications/Jurimetrics/winter-2024/teaching-law-in-the-age-of-generative-ai.pdf).

46 *Farber*, Harmonizing AI and human instruction in legal education: a case study from Israel on training future legal professionals (2024), International Journal of the Legal Profession, volume 31, issue 3, 2024, S.7, <https://ssrn.com/abstract=5043572>.

47 *Heetkamp/Brachtendorf*, DNH 3/2024, S. 14, [https://www.hlb.de/fileadmin/hlb-global/downloads/dnh/full/2024/DNH\\_2024-3.pdf](https://www.hlb.de/fileadmin/hlb-global/downloads/dnh/full/2024/DNH_2024-3.pdf).

48 Hierzu *Eder* (Fn. 19).

49 *Ogorek/Barthle* (Fn. 37), S. 117.

Vorlesung vorbereiten und die Präsenzzeit dann für die praktische Anwendung und vertiefte Diskussion reserviert bleibt.<sup>50</sup>

Ziel bei alledem muss sein, einerseits den Stoff zu erarbeiten und zu vertiefen, andererseits auch die kritische Auseinandersetzung mit KI-generierten Inhalten zu trainieren. Gleichzeitig stellen sich Herausforderungen beim Zeit- und Ressourcenmanagement. Denn die zusätzlich benötigte Zeit fehlt für die klassische Stoffvermittlung. Auch dürften nicht überall genügend geeignete Personal, Räume, Rechner und Budget vorhanden sein.

### 3. Neue Möglichkeiten für das Lernen

Auch beim Lernen könnte generative KI unterstützen. Hierzu gibt es folgende Ideen:

Chatbots könnten Lernen personalisiert, interaktiv und somit motivierend gestalten. Bereits in der Lehr- und Lernpraxis angekommen ist seit dem Sommersemester 2025 der „OneTutor“, der an der TU München entwickelt wurde.<sup>51</sup> Dieser ermöglicht es Studierenden, jederzeit Fragen zu Vorlesungsinhalten zu stellen und automatisch generierte, aber von Lehrenden kuratierte Quiz-Aufgaben zu lösen. Da nur die hinterlegten Vorlesungsmaterialien als Wissensbasis dienen<sup>52</sup>, wird die Gefahr von Halluzinationen reduziert. Wenn zu einer Fragestellung keine Informationen in den hinterlegten Materialien vorhanden sind, kommuniziert das System dies transparent.

Im Rahmen eines dreijährigen Forschungsprojekts wird der OneTutor an insgesamt zehn bayerischen Hochschulen und Universitäten in verschiedenen Fachrichtungen auf seine Wirksamkeit hin untersucht werden.<sup>53</sup> Im juristischen Bereich wird der Tutor etwa an der Ludwig-Maximilians-Universität München in mehreren Veranstaltungen erprobt. Erste Eindrücke deuten darauf hin, dass insbesondere die Möglichkeit, jederzeit ohne Hemmschwellen Fragen stellen zu können, für Studierende einen erheblichen Mehrwert darstellt. Zudem

macht es, wie der *Verfasser* selbst ausprobieren konnte, Spaß, mit dem OneTutor zu arbeiten.

Auch weitergehende Ideen sind denkbar. So wird etwa ein Chatbot diskutiert, der ausgehend von einem eingegebenen Sachverhalt durch die Prüfung der einzelnen Tatbestandsmerkmale der in Frage kommenden Straftaten führt.<sup>54</sup>

KI-gestützte Lernsysteme können Feedback zu Texten geben, etwa zur Einhaltung des Gutachtenstils.<sup>55</sup> Der Vorteil liegt in der sofortigen Rückmeldung, anstatt wie bislang bei schriftlichen Arbeiten geraume Zeit auf die Korrektur warten zu müssen. Die Qualität und Schnelligkeit des Feedbacks sowie ausreichende Übungsgelegenheiten spielen aus lernpsychologischer Sicht eine zentrale Rolle.<sup>56</sup>

Möglich werden „befragbare“ Gesetzestexte<sup>57</sup> sowie Lernplattformen, die das Lernverhalten und den Wissensstand der Lernenden kontinuierlich analysieren und darauf basierend maßgeschneiderte Lernressourcen und Übungsaufgaben anbieten.<sup>58</sup> Wenn man dies weiterdenkt, wären sogar persönliche KI-Tutoren denkbar, die als ständiger Lernbegleiter fungieren. Sie könnten Verständnisfragen beantworten und auf den aktuellen Wissensstand zugeschnittene Übungsaufgaben oder sogar -spiele (*gamification*) erstellen. Über längere Zeiträume können diese Systeme Stärken und Schwächen der Lernenden identifizieren und gezielt Wissenslücken schließen.<sup>59</sup>

### 4. Probleme und Fallstricke

Es gibt jedoch Argumente, die gegen den Einsatz von KI-Tools in der Ausbildung sprechen. Hier sind neben der bereits angesprochenen Fehleranfälligkeit und den Halluzinationen folgende Aspekte zu nennen:

a) Mögliche Einschränkung der Fähigkeit zu eigenständigem, kritischem Denken

Der unreflektierte Einsatz von KI birgt die Gefahr, dass Studierende und Referendare wesentliche Denkprozesse

50 Ogorek/Barthle (Fn. 39).

51 Vgl. <https://onetutor.ai/>

52 Vgl. zu RAG oben Fn. 9.

53 <https://www.bidt.digital/forschungsprojekt/effektivitaet-generativer-ki-tutoren-in-der-hochschullehre/>.

54 Farber (Fn. 46), S.4.

55 Ein Beispiel hierfür ist LegalWriter (<https://legal-writer.com/>) oder das Projekt DeepWrite an der Universität Passau, das auch zur Klausurkorrektur eingesetzt werden soll, <https://www.uni-passau.de/deepwrite>.

56 Vgl. Kahnemann, Schnelles Denken, langsames Denken, 16. Aufl., S. 298 ff.

57 Siehe zu einem solchen Ansatz <https://frag-das-gesetz.de/>.

58 In diese Richtung zielt z.B. das Angebot wie [dskrpt.de](https://www.dskrpt.de).

59 Vgl. Murray, Artificial Intelligence for Learning the Law: Generative AI for Academic Support in Law Schools and Universities, *Journal of Law & Technology at Texas*, 8 Tex. J. L. & Tech. (im Erscheinen, 2025), [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4946680](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4946680).

nicht mehr selbstständig durchlaufen und diese somit nicht mehr in Fleisch und Blut übergehen. Wie aktuelle Studien nahelegen, könnte das Auslagern von Denkprozessen an KI-Systeme, sog. *cognitive offloading*, die Fähigkeit zu kritischem Denken beeinträchtigen<sup>60</sup> oder gar zu einer Abhängigkeitsspirale führen. Hinzu kommt der *automation bias*, also die menschliche Tendenz, Vorschlägen einer Maschine übermäßig zu vertrauen.<sup>61</sup>

Diese Effekte sollten in der juristischen Ausbildung nicht unterschätzt werden. Kritisches Denken ist der *unique selling point* guter Juristen und sollte daher insbesondere auch gegenüber KI-Tools zur Anwendung kommen. Es ist daher wichtig, dass angehende Juristen lernen, KI nicht zur Auslagerung des Denkens, sondern zur Erlangung neuer Perspektiven und tieferen Verständnisses zu benutzen.<sup>62</sup>

#### b) Entmenschlichung der Ausbildung

Es besteht ferner die Gefahr, dass in einem Brei von KI-generierten Texten (sog. *AI slop*) der Fokus auf das Wesentliche und der sozial wichtige Austausch im Studium verloren geht. Ein auf die Spitze getriebenes Beispiel: Wenn KI-Tools Vorlesungsmaterialien und Prüfungsaufgaben erstellen, Studierende und Referendare beim Lösen selbiger Aufgaben unterstützen und diese Arbeiten anhand einer ebenfalls maßgeblich von KI erstellten Musterlösung bewerten, würde der Faktor Mensch weitgehend aus der Ausbildung eliminiert – ein beunruhigender Gedanke. Zudem würde eine allzeit bereitstehende KI möglicherweise die Hemmung weiter steigern, die Dozentin einfach direkt zu fragen.

#### c) Rechtliche Aspekte

Daneben sind auch handfeste datenschutz- und urheberrechtliche Aspekte zu beachten, wenn personenbezo-

gene Daten oder urheberrechtlich geschützte Werke in KI-Systeme eingegeben werden. Viele der marktgängigen KI-Systeme speichern Nutzereingaben zur Modellverbesserung. Aber auch wenn diese Funktion ausgeschaltet wird, kommt es immer noch zur Verarbeitung personenbezogener Daten (vgl. Art. 4 Nrn.1, 2 DSGVO) beziehungsweise zur Vervielfältigung urheberrechtlich geschützter Werke (vgl. §§ 15 Abs. 1 Nr. 1, 16 UrhG) mit entsprechenden rechtlichen Fragestellungen, auf die hier nicht weiter eingegangen werden kann.

#### d) Moralische Aspekte

Auch jenseits der Rechtmäßigkeit wirft der Entstehungsprozess der KI-Modelle Fragen auf. Häufig wird thematisiert, dass diese Modelle, da sie mit Daten der Vergangenheit trainiert wurden, bestehende Vorurteile im Rechtssystem reproduzieren und verstärken können (*bias*).<sup>63</sup>

Zudem wird häufig die Intransparenz der Entscheidungsprozesse von KI (black box) thematisiert.<sup>64</sup> Sogenannte Reasoning-Modelle<sup>65</sup>, die vermeintlich ihre Gedankengänge offenbaren, schaffen nur scheinbar Transparenz.<sup>66</sup>

Die fehlende Transparenz des Entstehungsprozesses der KI-Modelle ist sogar noch gravierender. Hier sind insbesondere die unklare Urheberrechtslage<sup>67</sup> hinsichtlich der Trainingsdaten, aber auch datenschutzrechtliche Fragestellungen<sup>68</sup> zu nennen. Zudem wird die „Persönlichkeit“ von KI-Systemen außer durch die Trainingsdaten auch durch ihr Design (insbesondere *Finetuning* und *Alignment*<sup>69</sup>) bestimmt. Hier ist völlig intransparent, welchen „Wertekompass“ die KI-Systeme von ihren Schöpfern mitbekommen.<sup>70</sup>

Diese Aspekte sollten zu einem Moment des Innehaltens führen, bevor solche Modelle angehende Juristen

60 Lee et al., The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers (2025), <https://www.microsoft.com/en-us/research/publication/the-impact-of-generative-ai-on-critical-thinking-self-reported-reductions-in-cognitive-effort-and-confidence-effects-from-a-survey-of-knowledge-workers/>; Gerlich, AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking (2025), *Societies* 2025, 15(1), 6, <https://doi.org/10.3390/soc15010006>.

61 Vgl. [https://en.wikipedia.org/wiki/Automation\\_bias](https://en.wikipedia.org/wiki/Automation_bias).

62 Zu den teils problematischen Anwendungstendenzen von Studierenden in MINT-Fächern siehe Anthropic Education Report: How University Students Use Claude (2025), insbesondere den Abschnitt "How students interact with AI", <https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude>.

63 Vgl. z.B. Goswami, Revolutionizing Legal Education: The Role of Artificial Intelligence in Shaping the Future of Law Teaching and Learning, S. 19f. (2025); [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5123719](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5123719); nichts anderes gilt aber im Prinzip für

die klassische Recherche von Rechtsprechung und Literatur in Datenbanken oder Kommentaren.

64 Vgl. z.B. Brockmann, *NdsVBl* 2023, 287 (294).

65 Vgl. oben Fn. 8.

66 vgl. Chen et al., Reasoning Models Don't Always Say What They Think (2025), [https://assets.anthropic.com/m/71876fabefoed4/original/reasoning\\_models\\_paper.pdf](https://assets.anthropic.com/m/71876fabefoed4/original/reasoning_models_paper.pdf).

67 Wagner, *MMR* 2024, 298; Dornis, *CR* 2024, 765, 831.

68 Hüger, *ZfDR* 2024, 263; Paal, *ZfDR* 2024, 129.

69 *Finetuning* bezeichnet die Anpassung eines KI-Modells an spezifische Aufgaben durch zusätzliches Training mit spezialisierten Datensätzen. *Alignment* beschreibt den Prozess, bei dem KI-Modelle so konfiguriert werden, dass ihre Ausgaben mit menschlichen Werten, Präferenzen und Normen übereinstimmen – etwa durch Bewertung ihrer Antworten durch menschliche Gutachter.

70 Ein Beispiel ist das chinesische KI-Modell „Deep Seek“, das keine Auskunft über bestimmte geschichtliche Ereignisse geben will. Das amerikanische KI-Modell Grok3 zensiert kritische Inhalte zu Elon Musk und Donald Trump.

systematisch beim Lernen und bei der Beurteilung von Rechtsfragen unterstützen.<sup>71</sup> Es wäre geradezu ironisch, wenn rechtswidrig trainierte KI-Modelle zur Ausbildung junger Juristen verwendet würden.

#### e) Zukünftige Kosten und Monopole

Derzeit sind generative KI-Systeme kostenlos oder zu niedrigen Preisen verfügbar. Es ist jedoch davon auszugehen, dass dieses Geschäftsmodell – da defizitär – nicht langfristig bestehen wird. Spezialisierte juristische KI-Anwendungen wie rechtsspezifische Recherche-Tools oder fachspezifische Sprachmodelle sind bereits heute mit erheblichen Lizenzkosten verbunden. Die Finanzierung solcher Systeme könnte sowohl für Universitäten als auch für Studierende eine erhebliche Zusatzbelastung darstellen. Hinzu kommen Aufwände für die Compliance nach DSGVO und ggf. auch KI-VO. Es scheint zudem empfehlenswert, Abhängigkeiten von monopolartigen, vorwiegend ausländischen Anbietern zu vermeiden. Eine mögliche Alternative stellen Open-Source-Ansätze dar.<sup>72</sup>

#### 5. Zwischenfazit

Der Einsatz von KI-Tools in Studium und Referendariat ist differenziert zu betrachten. Ein großer Vorzug der Tools ist sofortiges, individuelles Feedback, das im Universitäts- und Referendariatsalltag sonst nur im direkten mündlichen Austausch zu bekommen ist. Auch der Spaßfaktor sollte nicht geringgeschätzt werden. Ob KI-Tools die nötige inhaltliche Qualität aufweisen und ob sie den Lernprozess tatsächlich unterstützen können, hängt u.a. von den eingesetzten KI-Modellen, dem zur Verfügung stehenden Material, der Implementierung und einem bewussten, kritischen Umgang ab.<sup>73</sup> Die Tools sollten daher wissenschaftlich evaluiert werden, bevor sie auf breiter Front angewendet werden. Der OneTutor der TU München ist hier als mustergültiger Ansatz zu nennen.

Unabhängig davon sollte zu Beginn des Jurastudiums der Fokus darauf liegen, das juristische Handwerkszeug zu erlernen. Dazu gehört das Fördern von Hartnäckigkeit, Lust, den Dingen auf den Grund zu gehen, und die Fähigkeit, ein gutes Argument wertzuschätzen und ein

schlechtes als solches zu entlarven. Es gehört zum Lernprozess dazu, sich auch einmal an einem schwierigen Problem die Zähne auszubeißen. Der vorschnelle Rückgriff auf KI könnte zur Bequemlichkeit verführen und den Lernerfolg gefährden. Daher sollte der Einsatz solcher Tools gerade in der Anfangsphase besonders gut beachtet werden.<sup>74</sup>

#### IV. Prüfen „durch KI“?

Die breite Verfügbarkeit generativer KI-Systeme gibt auch Anlass, über veränderte Prüfungsformate nachzudenken. Handschriftliche Klausuren, bei denen papiergebundene Gesetzestexte und Kommentare zum Einsatz kommen, spiegeln die Anforderungen der modernen Rechtspraxis nicht wider.

##### 1. Digitalisierung als Grundvoraussetzung

Notwendig ist im ersten Schritt die Digitalisierung des Prüfungswesens. Diese vollzieht sich gegenwärtig in mehreren Stufen.

Die Bundesländer führen zunehmend das elektronische Staatsexamen (E-Examen) ein, wobei typischerweise zunächst das Zweite Staatsexamen und anschließend die Erste Juristische Staatsprüfung digitalisiert werden.<sup>75</sup> Die Studierenden verfassen ihre Klausuren dabei am Computer, greifen aber weiterhin auf analoge Gesetzestexte zurück. In der universitären Ausbildung werden Arbeiten vielfach noch mit Stift und Papier verfasst. Im Referendariat in Bayern können aber z.B. alle Übungsklausuren bereits auf dem eigenen Laptop geschrieben werden.

Der nächste Entwicklungsschritt betrifft die digitale Korrektur der Prüfungsarbeiten – hier erfolgt auch bei den E-Examina teils noch ein Ausdruck der zuvor digital geschriebenen Arbeiten und eine handschriftliche Korrektur. Im Referendariat in Bayern werden die Übungsklausuren – anders als im Staatsexamen – bereits elektronisch korrigiert.

Der dritte Schritt ist die Digitalisierung der Hilfsmittel. Dies wäre für Studierende auch deshalb eine Erleichterung, weil das bisher händische Nachsortieren von Gesetzestexten nebst Übertragen von Markierungen und

71 Hinzu kommen Nachhaltigkeitsfragen (Verbrauch großer Mengen von Energie, Wasser und Ressourcen) und fragwürdige Arbeitsbedingungen beim *Finetuning* und *Alignment* (s. oben Fn. 69), siehe TIME, “OpenAI Used Kenyan Workers on Less Than \$2 Per Hour to Make ChatGPT Less Toxic”, <https://time.com/6247678/openai-chatgpt-kenya-workers/>.

72 Ein Beispiel ist das KI-Assistenzsystem F 13 aus Baden-Württemberg für Verwaltungsmitarbeitende, [https://stm.baden-wuerttemberg.de/de/service/presse/meldung/pid/kuenstliche-intelligenz-](https://stm.baden-wuerttemberg.de/de/service/presse/meldung/pid/kuenstliche-intelligenz-in-der-verwaltung)

[in-der-verwaltung](https://stm.baden-wuerttemberg.de/de/service/presse/meldung/pid/kuenstliche-intelligenz-in-der-verwaltung).

73 So auch Schwarcz et al (Fn. 10), S. 60.

74 Noch strenger Choi/Monahan/Schwarcz (Fn. 11), S. 53; vgl. auch die differenzierende Darstellung bei Bliss (Fn. 43), S. 149f.

75 Vgl. Meldung bei beck-aktuell vom 25.3.2024: E-Examen: Staatsprüfung 2.0 oder Digitalisierung auf Deutsch?, <https://rsw.beck.de/aktuell/daily/meldung/detail/e-examen-staatspruefung-erstes-zweites-examen-laender-berlin-nrw>.

Kommentierungen entfällt.<sup>76</sup> Erste Universitäten machen Schritte hin zur einer solchen Voldigitalisierung von Klausuren.<sup>77</sup>

## 2. Welche Rolle kann KI in Prüfungen spielen?

Vor einer weitgehenden Digitalisierung des Prüfungswesens macht der Einsatz von KI nicht viel Sinn. Schon jetzt lohnt es sich aber, über Chancen und Risiken nachzudenken.

### a) Klassische Klausuren

Klassische Aufsichtsarbeiten ohne KI-Unterstützung bleiben die Basis, um grundlegende juristische Denk- und Argumentationsmuster abzuprüfen. Ohne dieses Fundament können auch KI-gestützte Werkzeuge nicht sinnvoll genutzt werden. Diese Prüfungsart sollte daher weiterhin eine wichtige Rolle spielen.<sup>78</sup>

### b) Haus- und Seminararbeiten

Die Rolle von klassischen Haus- und Seminararbeiten ist mittlerweile fraglich geworden.<sup>79</sup> Es ist davon auszugehen, dass KI-Modelle auch in deutscher Sprache und im deutschen Recht bald einen Leistungsstand erreichen, der häusliche Arbeiten zur Leistungsdifferenzierung zunehmend weniger geeignet macht.<sup>80</sup> Das Niveau der jeweils verfügbaren KI-Tools ist der allgemeine Mindeststandard, der ohne jede Eigenleistung erreichbar ist. Grundlegende Schwächen der menschlichen Bearbeiter würden überdeckt statt offengelegt.<sup>81</sup>

Nicht primär für die Ermittlung des individuellen Leistungsstandes, aber jedenfalls didaktisch sinnvoll bleiben häusliche Arbeiten dann, wenn der Einsatz von KI-Tools in einem Anhang zur Arbeit dokumentiert und reflektiert wird.<sup>82</sup>

Sicher ist, dass ein KI-Verbot bei der Anfertigung von Haus- und Seminararbeiten keinen Sinn ergibt. Denn ein Verstoß dagegen kann in der Regel nicht nachgewiesen, ein Verbot somit nicht durchgesetzt werden.<sup>83</sup>

Außerdem sollen die Arbeiten ja gerade dabei helfen, zu lernen, wie man später recherchiert und arbeitet – und das wird mit KI sein.

### c) Prüfungen, bei denen KI-Einsatz erlaubt ist

In fortgeschrittenen Studienphasen sowie im Referendariat könnten Klausuren angeboten werden, bei denen Einsatz von KI-Tools ausdrücklich erlaubt ist. Dafür spricht, dass dies der künftigen Arbeitsrealität der Absolventen entspricht. Es könnte abgeprüft werden, wie gut die Studierenden mit Tools umgehen können und welche Ergebnisse sie in einer gegebenen Zeitspanne damit erzielen. Um eine Leistungs nivellierung durch KI-Tools zu vermeiden, müssten Aufgabenstellungen anspruchsvoller werden und noch mehr als bisher die Transferleistung in den Vordergrund stellen.<sup>84</sup> Die Chancengleichheit erfordert es zudem, dass alle Studierenden Zugriff auf die gleichen Tools haben.<sup>85</sup> Zudem werden sich Studierende und Referendare für solche Formate wohl nur erwärmen können, wenn auch in der entscheidenden Abschlussprüfung der Einsatz von KI zugelassen ist. Bis dahin ist es allerdings noch ein weiter Weg.

Weniger aufwändig wären Prüfungsformate, in denen ein KI-generierter Text, z.B. ein Urteils- oder Schriftsatzentwurf, auf Fehler überprüft werden muss.<sup>86</sup> Denn auch dies wird künftig eine in der Praxis übliche Aufgabenstellung sein.

### d) Mündliche Prüfungen

Um das wahre Leistungsniveau sowie Kompetenzen wie sicheres Auftreten, Verhandlungsgeschick und mündliche Ausdrucksfähigkeit festzustellen, könnte an eine Ausweitung mündlicher Prüfungen gedacht werden. Allerdings sind diese mit einem hohen Aufwand verbunden und werfen mit Blick auf die Chancengleichheit Fragen eigener Art auf.

Eine mögliche Verbesserung wäre die bereits erwähnte KI-gestützte Simulation von Gerichtsverhand-

76 Ein Angebot hierfür macht <https://lexmea.de>.

77 An der Universität Bielefeld wurde im Dezember die erste komplett digitalisierte Juraklausur geschrieben, vgl. Meldung bei LTO vom 18.12.2024; „Die erste komplett digitalisierte Juraklausur Deutschlands“; <https://www.lto.de/karriere/jura-studium/stories/detail/klausur-digitalisierung-jurastudium-pilot-online-gesetzbuch>.

78 So auch Heckmann/Rachut, OdW 2024, 85 (91 ff).

79 So auch Brockmann (Fn. 64), 290; Schwartmann, MMR-Aktuell 2023, 455536.

80 Vgl. oben Fn. 21.

81 In diese Richtung auch Choi et al. (Fn. 11), S. 53.

82 Vgl. Mayrhofer (Fn. 42).

83 Vgl. die Entscheidung VG München NJW 2024, 1052 mit krit. Anmerkung Rachut zu einer versagten Zulassung zum Master-Studium aufgrund eines „verdächtig guten“ Essays; hierzu ferner Braegelmann, RD 2024, 188; Schwarz, RD 2024, 458; zum Ganzen Heckmann/Rachut (Fn. 78), S. 94 ff.

84 So auch Ogorek/Barthle (Fn. 39).

85 Zur Frage der Chancengleichheit ausführlich Heckmann/Rachut (Fn. 78), S. 92 Rn. 18; Ekardt/Klotz, MMR 2024, 545 (546) sowie Bliss (Fn. 45), S. 142 f.

86 Heckmann/Rachut (Fn. 78), S. 94.

lungen oder Mandantengesprächen. Dadurch wären jedenfalls die Rahmenbedingungen und Themen vergleichbarer und objektiver, vielleicht auch praxisnäher. Als Vorstufe könnte ein KI-Avatar eine mündliche Prüfung simulieren und so bei der Vorbereitung auf menschliche Prüfer helfen.<sup>87</sup>

#### e) Multiple-Choice

Es mag rückschrittlich klingen: Doch auch an die Einführung von Multiple-Choice-Tests könnte gedacht werden. Diese spiegeln zwar offensichtlich nicht die Anforderungen der Praxis wieder. Der Vorteil dieser Tests ist jedoch, dass sie ohne großen Aufwand automatisch, transparent und unter Wahrung der Chancengleichheit bewertet werden könnten. Hier könnte ein Blick auf die – als anspruchsvoll bekannte – Uniform Bar Examination weiterführen, bei dem Multiple-Choice-Fragen eines von mehreren Elementen ausmachen.<sup>88</sup>

### 3. Bewertung durch KI

Auch die KI-unterstützte oder maßgeblich von KI vorgenommene Korrektur rückt in den Bereich des Möglichen.<sup>89</sup> Dies würde Aufwand und Kosten reduzieren sowie ein schnelles, individuelles Feedback ermöglichen. Auch verspricht der Einsatz algorithmischer Systeme eine gegenüber der menschlichen Benotung<sup>90</sup> höhere Objektivität.

Andererseits hat sich in eigenen Experimenten des *Verfassers* mit anonymisierten Seminararbeiten im Vergleich zur menschlichen Korrektur eine Konvergenz zur Mitte gezeigt, da KI-Systeme grobe Fehler nicht konsequent ankreideten, andererseits originelle Gedanken nicht positiv genug würdigten. Alles in allem waren die Standard-KI-Systeme für die Korrektur nicht zu gebrauchen.

Bei klassischen Klausuren und Hausarbeiten, bei denen im KI-System eine Musterlösung hinterlegt ist, wird dies besser funktionieren. Jedoch erfordert die verbindliche Bewertung von Leistungen auch hier regelmäßig ein Mitdenken der Korrektorin, gerade wenn keine

Musterlösung existiert oder die Bearbeitung in vertretbarer Weise von der Musterlösung abweicht. Es erscheint fraglich, inwiefern eine KI solchen Gedankengängen folgen und diese fair bewerten könnte.<sup>91</sup> Hier fehlen, soweit ersichtlich, bisher empirische Erfahrungswerte. Ein Preprint legt jedoch nahe, dass eine KI-Korrektur jedenfalls bei kürzeren Aufgabenstellungen aus statistischer Sicht eine menschliche Korrektorin ersetzen könnte.<sup>92</sup>

Neben den genannten Qualitätsproblemen stellt sich hier die Frage der Akzeptanz. Eine Korrektur durch KI verspricht zwar Objektivität und Gleichbehandlung, andererseits stellt sich auch ein mulmiges Gefühl ein, wenn menschliche Gedanken durch Maschinen bewertet werden sollen.

Dieses mulmige Gefühl manifestiert sich in rechtlichen Grenzen: KI-Systeme, die bestimmungsgemäß für die Bewertung von Lernergebnissen verwendet werden sollen, sind grundsätzlich als Hochrisiko-KI im Sinne der KI-Verordnung anzusehen (Art. 6 Abs. 2 KI-VO i.V.m. Anhang III Nr. 3b). Dies bedeutet, dass solche Systeme zwar erlaubt sind, aber ab August 2026 strengen Regularien unterliegen, Art. 113 Abs. 1 KI-VO. Eine Ausnahme von der Risikoeinstufung würde nur gelten, wenn das KI-System das Ergebnis der Entscheidung über die Bewertung nicht wesentlich beeinflussen würde, Art. 6 Abs. 3 KI-VO, was aber bei einer Klausurbewertung durch das KI-System gerade der Fall ist. Somit ist bei KI-Korrektur-Software unter anderem ein Risikomanagementsystem (Art. 9 KI-VO) sowie eine Daten-Governance (Art. 10 KI-VO) einzurichten, menschliche Aufsicht sicherzustellen (Art. 14 KI-VO) und eine Grundrechte-Folgenabschätzung durchzuführen (Art. 27 KI-VO). Zudem sind, weil es bei Prüfungen stets auch um die Verarbeitung personenbezogener Daten geht, datenschutzrechtliche Vorgaben, insbesondere das grundsätzliche Verbot automatisierter Entscheidungen gemäß Art. 22 DSGVO, zu beachten.<sup>93</sup>

Darüber hinaus wirft der Einsatz von KI im Prüfungswesen auch wesentliche grundrechtliche Fragen

87 Vgl. *Jahn*, NJW-aktuell 18/2025, S. 19 zu einem an der Universität Bayreuth entwickelten KI-Avatar.

88 Siehe <https://www.ncbex.org/exams/ube/>; Teil der Prüfung ist die Multistate Bar Examination (MBE) als sechsstündiger Multiple-Choice-Test mit 200 Fragen.

89 Siehe etwa das Projekt Rechtsmentor von Prof. Thomas Hoeren <https://chatgpt.com/g/g-67c0248dd54c819191618b99577cc7c7-rechtsmentor-itm-gpt>.

90 Kritisch zur Objektivität juristischer Klausurbewertungen *Hufeld*,

ZDRW 2024, 59, <https://www.nomos-elibrary.de/de/10.5771/2196-7261-2024-1/zdrw-zeitschrift-fuer-didaktik-der-rechtswissenschaft-jahrgang-11-2024-heft-1>.

91 Etwas anderes würde für die automatische Korrektur von Multiple Choice Tests gelten. Diese könnten automatisch und basierend auf einfachen Regeln (also gerade nicht durch KI-Systeme) korrigiert werden.

92 *Fan et al.* (Fn. 17), S. 8 f.

93 *Heckmann/Rachut* (Fn. 78), S. 89f.

auf, weshalb es sich empfiehlt, sowohl die Erlaubnis oder Duldung als auch ein mehr oder weniger weitreichendes Verbot im Hochschulrecht zu verankern.<sup>94</sup>

## V. Zusammenfassung

Generative KI berührt das Selbstverständnis des juristischen Berufsstands insgesamt und damit auch das Selbstverständnis der juristischen Ausbildung.

Aus der vorstehenden Bestandsaufnahme lassen sich als zentrale Erkenntnisse festhalten:

- Generative KI wird den juristischen Arbeitsmarkt grundlegend verändern und dabei über „Routineaufgaben“ hinausgehen.
- Juristische Ausbildung muss weiterhin auf die Entwicklung kritischen und analytischen Denkens ausgerichtet sein.
- Sie muss zugleich die veränderten Anforderungen der juristischen Berufsbilder abbilden und daher verstärkt interdisziplinäre, digitale und KI-bezogene Kompetenzen integrieren – sowohl im Stoffplan („Lernen über KI“) als auch in Lehrformaten und Lerntools („Lernen mit KI“).
- Prüfungsformate sollten mit Blick auf die künftige

Arbeitsrealität weiterentwickelt werden, wobei die individuelle Leistung messbar bleiben muss.

- Bei allen Reformschritten sind neben den Chancen auch die vielfältigen Risiken von KI zu beachten.

Es ist Zeit, über den „Elefanten im Raum“ intensiv zu diskutieren. In die Diskussion sollten nicht nur Fakultäten, Justizprüfungsämter, Dozierende und Studierende einbezogen werden, sondern auch die juristische Praxis und alle Akteure, denen eine qualitativ hochwertige Ausbildung des juristischen Nachwuchses am Herzen liegt. Nur durch einen offenen Dialog und die Bereitschaft, auch bewährte Traditionen zu hinterfragen, kann die juristische Ausbildung zukunftsfähig gestaltet werden.

Dr. Sebastian Dötterl ist Richter am Oberlandesgericht München. Er ist Prüfer im Zweiten Juristischen Staatsexamen und als Lehrbeauftragter an der LMU München Mitveranstalter von Grundlagenseminaren zu den Themen Legal Tech, KI und Recht. Er sammelte zudem mehrjährige Erfahrung sowohl in der universitären Lehre als auch als nebenamtlicher Leiter von Referendar-Arbeitsgemeinschaften.

<sup>94</sup> Heckmann/Rachut (Fn. 78), S. 97f.