

Künstliche Intelligenz und „Fake Research Papers“ – strafrechtliche Aspekte

(Strafrechtliche Perspektiven auf ChatGPT & Co. in der
Wissenschaft)

Eric Hilgendorf

Gliederung

- Vorüberlegungen
- Strafrechtliche Einsatzszenarien
- Zusammenfassung
- Ausblick de lege ferenda

Ausgangslage

- „paper mills, which churn out bogus manuscripts containing text, data, and images partly or wholly plagiarized or fabricated, often massaged by ghost writers. Some papers are endorsed by unrigorous reviewers solicited by the authors . Such manuscripts threaten to corrupt the scientific literature, misleading readers and potentially distorting systematic reviews. The recent advent of artificial intelligence tools such as ChatGPT has amplified the concern.“
- 34 % der Forschungspapiere in den Neurowissenschaften gefälscht, 24 % in der Medizin? (Jahr 2020)
- (Science vom 12. May 2023, <https://www.science.org/content/article/fake-scientific-papers-are-alarmingly-common>. Ferner zu „paper mills“ Nature vom 25.3.2021, S. 516 ff.)

ChatGPT & Co. und neue Formen wissenschaftlichen Fehlverhaltens

- Neue, extrem leistungsstarke Technologie zum Erzeugen von Texten
- Vielfältige Einsatzmöglichkeiten auch in der Wissenschaft:
 - Sammeln und Zusammenfassung von Informationen zur Erstorientierung
 - Sprachliche Verbesserungen
 - Übersetzungen
 - ...
 - Automatisiertes „Ghostwriting“
- Grundproblem: Nicht selten ist die sozialetische/wissenschaftsethische Bewertung der neuen Möglichkeiten zweifelhaft.

Beispiele für den Einsatz von ChatGPT

- Themensuche
- Strukturieren eines Themas
- Suche nach Vorarbeiten
- Formulierung von Textvorschlägen
- Übernahme fremder Textpassagen/Bilder mit/ohne Veränderungen
 - Streichen einzelner Sätze
 - Austauschen von Worten
 - Veränderung des Stils
 - Übersetzung
 - ...
- Übernahme/Generierung ganzer Texte

Textformen

- Private Ideensammlung
- Exposé eines Antrags/einer Publikation
- Förderantrag
- Publikation
 - Geisteswissenschaftlicher Aufsatz
 - Darstellung medizinischer Forschungsergebnisse
- „Fake Research Papers“ als Herausforderung!

Verwandte Probleme

- „Abschreiben“/Plagiarismus
- „Ghostwriting“
- „Copy and Paste“
 - aber auch:
- Einsatz von Suchmaschinen
- Begriffsergänzung (Google usw.)
- Automatische Sprachfehlerkorrektur

Blick in die (sehr nahe) Zukunft

- KI generiert Text, Sprache und Video
- Beratungsgespräche mit KI; universitätsinterne Angebote?
- Auftrag an KI, relevante Forschungsthemen zu finden und Anträge zu generieren
- Automatisiertes Einreichen von (jeweils leicht modifizierten) Texten bei einer größeren Zahl von Zeitschriften
- KI-Detektionssoftware bei Verlagen
- Pflicht zur Verwendung von KI? KI-Verwendung zur Qualitätssicherung.
-

Strafrecht und neue technische Möglichkeiten

- Auftreten neuer Technologien (wie ChatGPT) schafft neue Handlungsmöglichkeiten, die teils positiv, teils negativ bewertet werden
- Nicht alles, was negativ bewertet wird, ist deswegen schon strafbar
- Strafrecht taugt nur sehr eingeschränkt als Quelle für die Bewertung neuer Technologien (kaum „sittenbildende Kraft“ des Strafrechts)
- Strafrecht schützt bestimmte Rechtsgüter. Bislang gibt es keinen Straftatbestand, der speziell die Integrität wissenschaftlichen Arbeitens schützt.

Strafrechtliche Einsatzszenarien

- Konkrete Gefährdung/Schädigung der Rechtsgüter Anderer
- „Wissenschaftlicher Betrug“ bei nicht offengelegter Verwendung von ChatGPT oder ähnlichen Programmen
- Beteiligung Dritter (Herausgeber wissenschaftlicher Journale, Veranstalter, Gutachter, etc.)
- Eigene Strafbarkeitsrisiken der für Publikationsorgane Verantwortlichen
- Strafrechtlicher Schutz bei ungerechtfertigten Vorwürfen

Konkrete Gefährdung/Schädigung der Rechtsgüter Anderer

- **Beispiel 1:** ad hoc erzeugte Pseudo-Studie mit erfundenen Daten zu den Wirkungen eines bestimmten Medikamentes/Impfstoffes
 - Körperverletzung, Tötung (jeweils einschließlich Versuch)
 - Betrug (z.B. bei Verkauf der angeblichen Forschungsergebnisse)
- **Beispiel 2:** ad hoc erzeugte Dokumentation über angebliches Fehlverhalten einer anderen Person
 - Beleidigungsdelikte

„Wissenschaftlicher Betrug“ bei nicht offengelegter Verwendung von ChatGPT oder ähnlichen Programmen

- **Beispiel:** Ein Aufsatz über „ChatGPT und Strafrecht“ wird komplett über ChatGPT erzeugt und unverändert vorgetragen
 - Urheberrechtsverletzungen (?)
 - Datenschutzverstöße (?)
 - Betrug, § 263 StGB setzt voraus Täuschungshandlung, Irrtumserregung, Vermögensverfügung, Vermögensschaden + subjektiver Tatbestand. Dürfte nur in Ausnahmefällen (Vortrag/Publikation gegen Honorar? Reisekostenerstattung) gegeben sein.
 - Kommerzielle Herstellung von „fake research papers“ in „paper mills“: § 129 („Bildung krimineller Vereinigungen“)?

Probleme des „Wissenschaftsbetruges“ mittels KI

- **Konkret**

- Wie ist KI-Unterstützung zu bewerten? (Themensuche, Informationssuche und-strukturierung, automatisierte Zusammenfassung älterer Forschungsergebnisse, Fehlerkorrektur, ...)?

- **Allgemein:**

- „Wissenschaftsbetrug“ ist grds. kein Vermögensdelikt und damit kein Betrug i.S.v. § 263 StGB
- Problem der Bestimmung wissenschaftlichen Fehlverhaltens (zeit- und kulturelle Bewertung)
- KI-Unterstützung vs. KI-Autorschaft

Typologie wissenschaftlichen Fehlverhaltens

- Fehlverhalten generierende Konflikte in Forschergruppen (Mängel in der Leitungsverwaltung, Zugang zu Forschungsmaterial und gruppeninterne Konflikte, Forschungsbehinderungen)
- Probleme der Autorschaft (Autor- und Urheberschaftsprobleme, Plagiate, Ideendiebstahl)
- Fehlverhalten im Umgang mit Forschungsdaten (Dokumentationsdefizite, Datenmanipulation und Datenfälschung, Rechte an und Missbrauch von Daten)
- Organisations- und verfahrensfehlerhaftes Verhalten (Falschangaben, Verschweigen von Interessenkonflikten, Befangenheiten, Kompetenzmissbrauch)
- Wissenschaftliches Fehlverhalten in Prüfung, Lehre und Betreuung

Nach: H. Schulze-Fielitz, in Löwer/Gärditz (Hrsg.), Wissenschaft und Ethik, 2012, 1 ff.

Beteiligung Dritter (Herausgeber wissenschaftlicher Journale, Veranstalter, Gutachter, etc.)

- **Beispiel:** Herausgeber X weiß, dass sein „Starwissenschaftler“ Y seine „Forschungsergebnisse“ regelmässig von ChatGPT erstellen lässt. Trotzdem lässt er ihn immer wieder in seiner Zeitschrift publizieren/lädt ihn immer wieder zu (bezahlpflichtigen) Veranstaltungen ein, um weiterhin kommerziell erfolgreich zu sein.
 - Beihilfe und sogar Mittäterschaft
 - von Herausgebern, Veranstaltern, Gutachtern usw.
 - in Bezug auf von Autoren begangene Straftaten möglich,
 - aber wohl extrem selten
 - Kein doloses Zusammenwirken

Eigene Strafbarkeitsrisiken der für Publikationsorgane/Veranstaltungen Verantwortlichen

- **Beispiel 1:** wie letzte Folie
 - Betrug an (zahlenden) Abonnenten/Zuhörern?
- **Beispiel 2:** Herausgeber X ist verantwortlich dafür, dass eingereichte Texte mittels ChatGPT korrigiert/“aufgehübscht“ werden
 - Betrug
 - Datenschutzverstöße durch Nutzung der KI? (X wird wohl „verantwortliche Stelle“ i.S.d. Datenschutzrechts, wenn er selbst die Nutzung von ChatGPT veranlasst und kontrolliert)

Strafrechtlicher Schutz bei ungerechtfertigten Vorwürfen

- **Beispiel:** X wirft dem Y die unzulässige Nutzung von ChatGPT bei der Herstellung eines Beitrags vor
 - Beleidigungsdelikte, insbes. § 186 StGB
 - Falsche Verdächtigung, § 164 StGB
 - Vortäuschen einer Straftat, § 145d StGB
- **Verfahrensrechtliches Problem:** Öffentliche Vorwürfe wissenschaftlichen Fehlverhaltens sind häufig unverhältnismäßig, weil
 - die negativen Wirkungen der Publizität die Wirkung zivilrechtlicher oder strafrechtlicher Sanktionen (wenn sie denn gerechtfertigt wären) oft weit übersteigen
 - Außerkraftsetzung der Unschuldsvermutung und anderer Grundsätze des Strafverfahrens (in dubio, Verpflichtung auf Wahrheitserforschung); keine wirksamen Möglichkeit der Gegenwehr, insbes. bei anonymen Vorwürfen, fehlende rechtstaatliche Kontrolle des Verfahrens
 - Problemverschärfung durch Einsatz von KI-betriebenen „Denunziations-bots“?

Zusammenfassung

- Neue technikgenerierte Handlungsmöglichkeiten schaffen neue Möglichkeiten sozialschädlichen/als sozialschädlich empfundenen Verhaltens
- „Fake Research Papers“ als technische Fortentwicklung des Ghostwriting und anderer Formen der Nutzung verbotener Hilfsmittel
- Keine besondere „sittenbildende Kraft“ des Strafrechts im Kontext wissenschaftlichen Arbeitens
- Strafbarkeitsrisiken bei Nutzung von KI zur Erzeugung von Forschungsergebnissen/Texten gering, aber (v.a. im Kontext „bezahlter Wissenschaft“) durchaus relevant

Ausblick: Neuer Straftatbestand?

- Traditionelle Antwort:
 - Ausgangslage: Meinungs- und Wissenschaftsfreiheit
 - Eingriffe müssen dem Verhältnismäßigkeitsgrundsatz genügen
 - Neue Straftatbestände mögen grds. **legitim** und auch **geeignet** sein, um wiss. Fehlverhalten gegenzusteuern,
 - dürften jedoch nach bisher überwiegender Ansicht **nicht erforderlich** sein, da es eine Vielzahl anderer effektiver Präventions- und Sanktionsmöglichkeiten gibt.

Ausblick: Neuer Straftatbestand?

- Aber:
 - Wissenschaftsfreiheit schützt nicht die vorsätzliche Erstellung und Verbreitung von „fake research papers“
 - Was bedeutet es, wenn in „paper mills“ im großen Stil Fälschungen mittels KI erzeugt werden und ein Großteil der gefälschten Papiere aus dem Ausland kommt?
 - Integrität wissenschaftlicher Kommunikation als Voraussetzung des wissenschaftlichen Fortschritts (z.B. in der Medizin, Naturwissenschaften, Technikwissenschaften) als neues Rechtsgut?
 - Aufgabe: **Klärung der faktischen Problemsituation und Fortbildung der Standards korrekten wissenschaftlichen Verhaltens**

- Vielen Dank für die Aufmerksamkeit!