

Welche Software rettet die Welt? Ein Werkstattbericht*

Armin Glatzmeier, CeDiS, FU Berlin

1 Begrüßung und Einführung

Meine sehr geehrten Damen und Herren, liebe Kolleginnen und Kollegen,

welche Software rettet die Welt? Diese Frage ist ebenso präventiv wie rhetorisch und ich könnte an dieser Stelle mit der schlichten Antwort „keine“ enden. Damit wäre nun der verfügbare Zeitrahmen allerdings erheblich unterschritten, sodass ich zunächst meine Frage ein wenig spezifizieren möchte: Wenn ich über *die* Welt spreche, dann meine ich natürlich den überschaubaren Bereich der akademischen Welt und auch davon nur den Teilbereich der schriftlich verfassten Beiträge. Diese, so lautet die implizite Prämisse, bedürfen der Rettung. Zu klären bleibt also noch: wovon und mit welchen Mitteln?

FOLIE 2

Lassen Sie es mich in meine eigenen Worte kleiden: Ein Gespenst geht um in Europa – streng genommen nicht nur in Europa – das Gespenst des Plagiarismus. Alle Mächte der alten Akademie haben sich zu einer heiligen Hetzjagd gegen dieses Gespenst verbündet, Ombudsleute und Prüfungskommissionen, Softwareschmieden und Reviewer, französische Journale und deutsche Publizisten.

Sie haben meine kleine Anleihe sicher erkannt, ohne sich dabei von einer Software unterstützen zu lassen. Allerdings haben auch Sie sich dabei einer Heuristik bedient, eines Algorithmus, der das gehörte Wort unter anderem syntaktisch und semantisch mit verfügbaren Informationen vergleicht. Warum also diese Aufgabe nicht automatisieren?

Diese Forderung ist vielleicht auch deshalb so populär, weil sie nach *cutting edge* und *state of the art* klingt, weil sie Arbeitserleichterung verspricht und unbestechlich präzise Ergebnisse. In den Gesprächen, die wir im Rahmen unseres Projektes hausintern geführt haben, spiegeln sich diese Hoffnungen und Erwartungen ebenso wie Skepsis und die Frage nach dem eigenen Anspruch als Betreuer*in. Eine Frage war jedoch in fast allen Gesprächen zentral: Was kann sog. Anti-Plagiatssoftware realistisch leisten?

*Vortrag im Rahmen der uni.digital – teaching assesment, learning gehalten an der FU Berlin am 5. April 2019.
Es gilt das gesprochene Wort.

Im Rahmen unseres Projekts „Softwaregestützte Qualitätssicherung wissenschaftlicher Arbeiten“ haben wir genau diese Frage in einem explorativen Testsetting an verschiedene Softwarelösungen zur Erkennung von Textübereinstimmungen gestellt.

Ein kurzes Wort zur Begrifflichkeit: Ich werde im Folgenden solche Software als Anti-Plagiatssoftware – kurz APS – bezeichnen, obwohl diese Bezeichnung in zweierlei Hinsicht irreführend ist: Zum einen erkennt Software keine Plagiate, sondern lediglich übereinstimmende Zeichengruppen; zum anderen suggeriert der Begriff, dass Plagiate durch den Einsatz einer solchen Software bekämpft oder verhindert würden. Auch dies ist irreführend, sofern man eine grundsätzliche generalpräventive Wirkung vernachlässigt, da sie von einer Reihe intervenierender Variablen abhängt.¹ Dennoch hat sich der Begriff mittlerweile etabliert und ist in der sprachlichen Anwendung einfacher als die präzisere Beschreibung.

FOLIE 3

2 Projektauftrag

Der Auftrag unseres Projektes gliedert sich in drei Arbeitsbereiche. Das sind

- ⑩ zum ersten der Test und ggf. die Empfehlung von Anti-Plagiatssoftware für den universitätsweiten Einsatz;
- ⑩ zum zweiten die Evaluation strategischer, organisatorischer, technischer und rechtlicher Grundlagen und Rahmenbedingungen eines APS-Einsatzes;
- ⑩ und zum dritten die Identifikation, Planung und Konzeption zentraler und/oder dezentraler Unterstützungsstrukturen zur Plagiatsprävention bzw. zur Sicherung guter wissenschaftlicher Praxis.

Von diesen Aufgaben ist die erste abgeschlossen. Mit der Zweiten haben wir begonnen. Die Dritte wird auf den Ergebnissen der vorangehenden aufbauen. Der Schwerpunkt meines Vortrages wird somit auf dem Softwaretest liegen. Auf weiterreichende Fragen können wir gerne in der Diskussion eingehen.

FOLIE 4

3 Softwaretest

3.1 Vorbemerkung

In unserem Projekt befassen wir uns mit Softwarelösungen für unterschiedliche Einsatzgebiete: Dies ist zum einen der Vergleich von Textidentitäten einer schriftlichen Prüfungslei-

¹Z. B. Wie konsequent wird in Plagiatsfällen verfahren, handelt es sich um eine intendierte oder unintendierte Fremdtextübernahme usw.

tung mit bereits andernorts veröffentlichten oder öffentlich zugänglichen Texten durch Abgleich mit Internetquellen. Zum anderen geht es um den Vergleich schriftlicher Prüfungsleistungen mit unveröffentlichten Prüfungsleistungen anderer Studierender („Querverprobung“).² In beiden Fällen sind hinsichtlich der zu prüfenden Textarten zwei grundlegend verschiedene Einsatzszenarien möglich: (a) zum einen der Abgleich wissenschaftlicher Abhandlungen sowie (b) zum anderen im Bereich der Informatik, der Vergleich von Programmcode, die unterschiedlich adressiert werden müssen. Die folgenden Ausführungen beziehen sich auf kommerzielle Softwarelösungen für den Abgleich wissenschaftlicher Abhandlungen mit freien und kommerziellen digitalisierten Quellen.

Zur Klärung dieses Sachverhalts haben wir zwei unterschiedliche Herangehensweisen gewählt. Zum einen baten wir Fachbereiche, Zentralinstitute und Lehrende um die Übersendung von Arbeiten, bei denen nach Möglichkeit ein substantiierter Anfangsverdacht dokumentiert wurde. Zum anderen testeten wir die verschiedenen Produkte unseres Portfolios mit zwei standardisierten Testdokumenten unter der Fragestellung, wie viele der in den Dateien eingebrachten Einzelplagiate von den unterschiedlichen Softwarelösungen erkannt werden. Dieser zweite Zweig unserer Testreihe ist Gegenstand dieses Vortrags.

3.2 Erkenntnisinteresse

Methodisch nahm das Testverfahren eine Umkehrung der Ausgangshypothese vor, auf der die Prüfung eines Textes auf potentielle Plagiate normalerweise beruht.³ Da bekannt war, dass die Testdokumente zu 100 Prozent aus Plagiaten bestanden und die Einzelplagiate in den Texten vorab definiert und beschrieben wurden, stellte sich nicht die Frage, ob der zu prüfende Text Plagiate enthält.

FOLIE 5

Unsere Hauptfrage zielte vielmehr darauf,

- ⑩ wie viele der Plagiate unserer Testdokumente und welche Formen von Plagiaten erkannt werden.

Als Nebenbefunde erhofften wir uns Erkenntnisse darüber,

- ⑩ ob die Erkennungsrate für verschiedene Dateiformate (odt, pdf und docx) der Testdokumente konstant ausfällt;
- ⑩ ob die Erkennungsrate von der Zugänglichkeit⁴ und dem (Datei-)Format der Quellen⁵ beeinflusst wird;

²Die beiden Varianten stellen zudem unterschiedliche rechtliche Herausforderungen dar.

³Aus Gründen der Fairness und der „Unschuldsvermutung“ ist die Ausgangsvermutung, dass der zu prüfende Text eine eigenständige Leistung ist. Die Entdeckung von plagiatverdächtigen Stellen zieht diese Nullhypothese in Zweifel. Sie kann aber erst mit dem harten Nachweis konkreter Plagiate zugunsten der Alternativhypothese („Die Arbeit enthält Plagiate“) aufgegeben werden.

⁴Frei oder kostenpflichtig.

⁵Print oder digital (und hier verschiedene Dateiformate).

- ⑩ ob die als identisch referenzierten Internetquellen den tatsächlich verwendeten Quellen entsprechen und
- ⑩ wie die Qualität möglicher alternativer Referenzstellen zu beurteilen ist.

Da unser primäres Erkenntnisinteresse darauf zielte, die Erkennungsleistung der verschiedenen Softwarelösungen vergleichend einschätzen zu können, verzichteten wir im Rahmen dieses Tests auf die Prüfung, ob die Texterkennung durch sogenannte Homoglyphe⁶ umgangen werden kann.

FOLIE 6

3.3 Softwareauswahl

Die Auswahl der Software beruht auf drei Auswahlritten: einer Marktevaluation, Recherchen zur Situation an vergleichbaren Hochschulen in In- und Ausland sowie der Auswertung bereits vorliegender Softwaretest. Im Ergebnis umfasste das Testportfolio anfangs folgende Produkte: CopyLeaks, Docoloc, iThenticate, PlagAware, PlagiarismChecker X, Plagius, PlagScan, SafeAssign, Turnitin, Unicheck und Urkund. In der ersten Phase des standardisierten Tests konnten somit neun internetbasierte Dienste und zwei Desktopapplikationen getestet werden. In der zweiten Testphase reduzierte sich die Zahl auf zehn Produkte, da die Testlizenz einer Desktopanwendung nach der ersten Feedbackrunde nicht erneuert wurde. Neben den kommerziellen APS versuchten wir die Plagiate mit den beiden Suchdiensten Bing und Google zu rekonstruieren. Da Google in der ersten Testphase deutlich besser abschnitt, wurde die manuelle Überprüfung mit Bing in der zweiten Testphase nicht fortgeführt.

An dieser Stelle kurz zu den Anbietern: Nach einer kurzen Vorstellung unseres Projekts erhielten wir für die vorab ausgewählten Produkte Testlizenzen bzw. Testzugänge. Diese konnten wir im Wesentlichen frei nutzen. Da die Bereitstellung der Testlizenzen zum Teil unter der Bedingung erfolgte, dass die Ergebnisse nicht öffentlich dokumentiert würden, sind die verschiedenen APS in den folgenden Diagrammen pseudonymisiert.

Im Gegenzug für die Testlizenzen stellten wir den Herstellern folgende Informationen zur Verfügung:

- ⑩ das Testdokument, in den beim Test verwendeten Dateiformaten,
- ⑩ eine synoptische Übersicht der im Testdokument verwendeten Plagiate mit exaktem Quellennachweis,
- ⑩ den bzw. die Prüfberichte ihres Produkts zu den eingereichten Testdateien,
- ⑩ eine Auswertung, wie ihr Produkt bezüglich der verschiedenen Plagiatsinstanzen im Testdokument abschnitt und

⁶„Plagiat“ ist eine homoglyphe Darstellung des Wortes „Plagiat“ mit überwiegend kyrillischen Zeichen. Häufig – aber nicht zwingend – fallen solche Umwandlungen durch ein abweichendes Schriftbild auf. Kürzere Ausdrücke lassen sich z. B. auf <http://homoglyphen.de/> (Abruf: 11.12.2018) umwandeln.

- ⑩ eine anonymisierte Vergleichsübersicht, aus der hervorgeht, wie ihre APS bei der Detektion der einzelnen Plagiatsinstanzen im Gesamtfeld abschnitt.

Besonders auffällige oder unerwartete Testergebnisse wurden dokumentiert und den Herstellern zur Kommentierung vorgelegt.

FOLIE 7

3.4 Setting des standardisierten Tests

Wie schon erwähnt, erfolgte der standardisierte Softwaretest in zwei Phasen. Für jede Testphase wurde ein Testdokument (Testdokument 1 und Testdokument 2) erstellt und in verschiedenen Dateiformaten gespeichert. In der ersten Testphase nutzen wir die Formate .odt und .pdf. In der zweiten Testphase kam das Format .docx hinzu. Alle APS des Portfolios wurden mit diesen Dateien getestet. Der Test mit Testdokument 1 erfolgte zwischen dem 5. und 27. September 2018. Testdokument 2 wurde am 21. November 2018 zeitgleich zur Prüfung gestellt. Beide Tests wurden mit den werksseitigen Standardeinstellungen durchgeführt.

Beide Testdokumente bestanden vollständig aus plagiierten Textübernahmen. Art und Umfang der einzelnen Plagiatsinstanzen wurden mit exaktem Quellennachweis und unserer Ausgangsvermutung, ob die jeweilige Plagiatsinstanz erkannt oder nicht erkannt werden würde, vor der Prüfung in einer Synopse dokumentiert.

Testdokument 1 wurde am 5. September 2018 als .odt in LibreOffice erstellt und mit Standardeinstellungen als .pdf exportiert. Bei der Erstellung wurden 24 Quellen genutzt – neun kommerzielle und elf freie Internetquellen sowie 4 Printpublikationen. Insgesamt enthält es 27 Plagiatsinstanzen, von denen für 23 die Ausgangsvermutung war, dass sie bei einer Softwareprüfung entdeckt werden würden.

Testdokument 2 wurde am 21. November 2018 als .odt in LibreOffice erstellt und mit Standardeinstellungen als .pdf und .docx exportiert. Bei der Erstellung wurden 29 Quellen genutzt – 21 kommerzielle und acht freie Internetquellen. Insgesamt enthält es 37 Plagiatsinstanzen, von denen für 33 die Ausgangsvermutung war, dass sie bei einer Softwareprüfung entdeckt würden.

Insgesamt deckten wir 23 Fächer mit Schwerpunkt auf den Geistes- und Sozialwissenschaften und fünf Sprachen – Deutsch, Englisch, Französisch, Italienisch und Spanisch – in den beiden Testdokumenten ab.

3.5 Befunde

Für die Bewertung der durch die APS gefundenen möglichen Referenzstellen erarbeiteten wir vor der Auswertung der Prüfberichte folgende Kategorien:

FOLIE 8

- ⑩ Als ‚Direkte Quelle‘ beschreiben wir Funde, bei denen die APS die unmittelbar bei der Erstellung des Testdokuments benutzte Quelle oder ein auflagengleiches Digitalisat referenziert.
- ⑩ Als ‚Indirekte Quelle (Typ 1)‘ beschreiben wir Funde, bei denen die APS bei *Büchern* eine andere Auflage bzw. Ausgabe, bei *Aufsätzen* eine identische oder teilidentische Zweitpublikation oder bei *Internetquellen* einen legalen Spiegel meldet. Wir betrachten diese Treffer als direkten Quellen gleichwertig.
- ⑩ Als ‚Indirekte Quelle (Typ 2)‘ beschreiben wir Funde sekundärer Quellen, aus denen die benutzte Primärquelle eindeutig erkennbar ist.
- ⑩ Als ‚Irreguläre Quelle‘ beschreiben wir Funde aus Internetforen oder Upload-Portalen, deren Inhalte von Usern – teils illegal – geteilt werden und die daher nicht als permanente Quellen betrachtet werden können. Diese wurden bei der Auswertung gesondert mit (+n) ausgewiesen (vgl. Tabelle).

Um dies zu verdeutlichen habe ich Ihnen einen Auszug aus der tabellarischen Auswertung mitgebracht, auf dem Sie kurz sehen können, wie wir bei dieser Auswertung vorgegangen sind.

FOLIE 9

Bei den folgenden Auswertungen, die ich Ihnen in Form einiger Diagramme vorstellen möchte, sind die Funde aus irregulären Quellen berücksichtigt, da es für die Dokumentation von Plagiaten inhaltlich unerheblich ist, wie die Quelle gefunden wurde. Für die juristische Bewertung einer Verwertbarkeit solcher Fundstellen mag dies anders sein.

FOLIE 10

3.5.1 Allgemeine Ergebnisse

Beginnen möchte ich mit einigen allgemeinen Ergebnissen zur Erkennungsrate der getesteten APS. Sie sehen hier zunächst die Prozentangaben aus den Prüfberichten der verschiedenen APS zu unserem Testdokument 1, die den Gesamtwert einer möglichen Fremdtextübernahme kennzeichnen. Für Google und Bing entfällt diese Angabe.

FOLIE 11

Vergleichen wir diesen automatisch generierten Wert mit dem effektiv auf Grundlage der Berichte als Plagiat erkennbaren Zeilenumfang in Testdokument 1, so fällt insbesondere auf, dass dieser in den meisten Fällen ein wenig höher ausfällt – eine Ausnahme bildet hier die APS 6. Zudem wird deutlich, dass die manuelle Suche nach Plagiaten mittels Google und Bing deutlich bessere Resultate zeitigt, als die Mehrzahl der getesteten APS.

Wie sich ebenfalls bereits im ersten Test zeigte, spielt für die maschinelle Erkennung von Plagiaten auch das Dateiformat eine Rolle, in dem der zu prüfende Text vorliegt. Wie die

Rücksprache mit den Herstellern ergab, lassen sich keine generellen Aussagen treffen, welches Dateiformat insgesamt die besseren Ergebnisse liefert, da die Encodierung des Textes in die verschiedenen Containerformate von unterschiedlichen – teils auch nutzerseitig variierbaren technischen Rahmenbedingungen abhängt.

Diese beiden Grundtendenzen – höherer Anteil effektiver Funde und Abhängigkeit der Erkennung vom Dateiformat – zeigen sich auch für Testdokument 2. Die im Vergleich zu Testdokument 1 insgesamt höhere Erkennungsrate führen wir auf die Erhöhung des Anteils der Copy&Paste-Plagiate zurück.

FOLIE 12

Hier sehen Sie wiederum zunächst die in den Prüfberichten der getesteten APS ausgewiesenen Prozentangaben, die eine mögliche Übereinstimmung mit anderen Quellen indizieren.

Und im Vergleich den prozentualen Zeilenanteil, der sich auf Grundlage der jeweiligen Funde als Plagiat rekonstruieren ließ

FOLIE 13

Was die Erkennung der verschiedenen Plagiatsformen, die wir in den Testdokumenten verarbeitet haben anbelangt, zeigt sich folgendes Bild.

FOLIE 14

Die Grafik zeigt die absolute Anzahl der verwendeten Plagiatsformen – das ist der blaue Balken, linksseitig skaliert – und den kumulierten Anteil der realisierten Funde an der Gesamtzahl aller möglichen Funde – hier der rechtsseitig skalierte rote Balken.

In Testdokument 1 verwendeten wir ein Strukturplagiat, d. h. wir haben die Gliederung eines Aufsatzes als Rahmen übernommen, in den wir die anderen Plagiate eingefügt haben. Zudem bauten wir vier Übersetzungsplagiate ein – dreimal Englisch nach Deutsch, einmal Französisch nach Deutsch. Der überwiegende Anteil bestand mit 19 Plagiaten aus einfachem Copy&Paste. Hinzu kamen sieben Paraphrasen und ein Bauernopfer, d. h. eine Plagiatsstelle, an der ein korrekt ausgewiesenes Zitat die großflächigere paraphrasierte Fremdtextübernahme verschleiern sollte.

Wie sich im Kern zeigt: Weder das Strukturplagiat noch die Übersetzungsplagiate konnten durch APS nachgewiesen werden. Die Erkennung von Copy&Paste – und hierzu können Sie auch die Bauernopfer-Referenz rechnen, da nur das darin enthaltene Zitat gefunden wurde – gelingt den APS im Schnitt deutlich besser als die Erkennung überarbeiteter Textpassagen.

Die gegenüber paraphrasierten Textübernahmen insgesamt höherer Erkennbarkeit einfacher Copy&Paste-Plagiate zeigt sich auch für Testdokument 2.

FOLIE 15

Die bessere Erkennung eines der beiden Strukturplagiate erklärt sich dadurch, dass einer der verwendeten Aufsätze zu den Klassikern der psychologischen Literatur zählt und sehr signifikante Überschriften aufweist.

Betrachtet man die kumulierten Funde aller APS im Test in Abhängigkeit von der Art der verwendeten Quellen, so ergibt sich für Testdokument 1 folgendes Bild.

FOLIE 16

Wie Sie sehen, werden frei zugängliche Internetquellen grundsätzlich besser erkannt als kommerzielle Inhalte. Überrascht hat uns die gute Erkennung der genutzten Printquellen, die sich im Gespräch mit den Herstellern dadurch erklärte, dass bei einigen APS auch die Ergebnisse von Internetsuchdiensten zum Abgleich einbezogen werden und die meist älteren Printquellen mittlerweile oft als Digitalisate vorliegen.

Die bessere Erkennung frei zugänglicher Quellen dokumentiert sich auch in der Auswertung unseres zweiten Testdurchlaufs.

FOLIE 17

Zum Vergleich möchte ich Ihnen zum Abschluss unserer Testergebnisse die Resultate der manuellen Rekonstruktion der verschiedenen Plagiate mittels Google für Testdokument 2 präsentieren.

FOLIE 18

Wie Sie sehen, konnten mit den verschiedenen Google-Suchdiensten die meisten Plagiatsstellen rekonstruiert werden.

Uns ist völlig klar, dass dieses Ergebnis mit Vorsicht zu genießen ist, zumal wir die Einzelplagiate und die Quellen kannten. Dies galt aber auch bezüglich der Rekonstruktion der Plagiate auf Grundlage der APS-Prüfberichte. Um das Risiko eines Bias zu minimieren, haben wir bei der Suche mit Google ein Vorgehen gewählt, das sich bei der manuellen Plagiatssuche in Originalarbeiten vielfach bewährt hat: Wir gaben an der Suchmaske nur einige sinntragende Begriffe aus dem zu prüfenden Text ein. Gleichzeitig hielten wir die Zahl der gewählten Suchbegriffe möglichst klein.

Was dieses Ergebnis bei aller Vorsicht deutlich macht, ist jedoch, dass mit maschineller Erkennung grundsätzlich eine bessere Abdeckung möglich wäre, als sich in unserem Vergleich für die getesteten APS gegenwärtig zeigt.

Ein kurzer Hinweis zur Methodik – auch um Fragen zur statistischen Validität und Generalisierbarkeit der Ergebnisse vorweg zu nehmen: Uns ist der rein explorative Charakter unserer Befunde bewusst (Stichwort: kleine Grundgesamtheit). Die Grundtendenz, die sich in den Ergebnissen abzeichnet, halten wir beim aktuellen Stand der Technik allerdings für generalisier- und reproduzierbar.

FOLIE 19

Dies betrifft die Befunde,

- ⑩ dass eine vollständige Erkennung aller Plagiate in einem Text mittels Softwareinsatz nicht geleistet werden kann;
- ⑩ dass frei zugängliche Netzinhalte besser erkannt werden als kommerzielle Angebote;
- ⑩ dass selbst wortlautgleiche Kopien freier Netzinhalte jedoch nicht immer zuverlässig erkannt werden;
- ⑩ dass paraphrasierte Übernahmen von Fremdtext schlechter erkannt werden als wortlautgleiche Übernahmen;
- ⑩ dass Übersetzungsplagiate generell nicht erkannt werden;
- ⑩ dass sinnentstellende Übernahmen fremder Texte und Gedanken nicht erkannt werden;
- ⑩ dass die Übernahme der Struktur eines Fremdtexes nicht erkannt wird.

Hinzu kommt ferner, dass mathematische Formeln für alle getesteten APS problematisch waren. Eine Erkennung plagiierten Bildmaterials ist mit keiner der kommerziellen APS möglich.

Wie sich im Test mit Originalarbeiten zeigte, kann gegenwärtig nicht erwartet werden, dass plagiatsbehaftete Arbeiten zuverlässig als solche erkannt werden. So hatten wir zwei Fälle, in denen Plagiate manuell nachgewiesen werden konnten, bei einer Softwareprüfung jedoch unentdeckt blieben.

Soweit die Negativliste.

Was unsere Befunde deutlich zeigen, ist, dass jede Software nur als Werkzeug bei der Prüfung auf Textplagiate betrachtet werden sollte. Insbesondere gilt, dass die maschinell erstellten Prüfberichte grundsätzlich einer Sichtung unterzogen werden müssen – wie auch das Verwaltungsgericht Baden-Württemberg 2015 feststellte.⁷

In der Praxis dürfte der Umstand, dass keine der getesteten APS sämtliche Plagiate erkannte, insofern unerheblich sein, als für die Bewertung einer Prüfungsleistung nicht der Gesamtumfang an Plagiaten nachgewiesen werden muss.

Was den Wunsch nach einer Arbeitserleichterung anbelangt, so zeigte sich, dass auch bei bekannter Plagiats- und Quellenlage die Dokumentation mit Internetsuchmaschinen einen erheblichen Mehraufwand bedeutet.

Insgesamt spricht somit aus dieser Perspektive einiges dafür, die Sucharbeit – nicht die Verantwortung – an APS zu delegieren.

Ob dieser Weg beschritten werden soll, hängt indes noch von einigen weiteren Fragen ab, die heute nicht Gegenstand meines Vortrages sind: der Klärung grundsätzlicher rechtlicher, organisatorischer und auch ethischer Fragen.

Was etwa, wenn der Plagiierte nur allzu bereitwillig mit der Übernahme seiner Ergebnisse einverstanden ist? Wenn Ghostwriter gerade darin ihren Lebensunterhalt finden, fremden

⁷VGH BW 9 S 327/14

Menschen zu wissenschaftlichen Meriten zu verhelfen? Wenn ein Unternehmen beispielsweise Pestizide herstellt, deren Vertrieb von staatlicher Zustimmung abhängig ist, und nun seine Auftragsstudien in den amtlichen Dokumenten übernommen findet?

Plagiate sind keine Kavaliersdelikte, auch wenn ihr Schaden meist auf den akademischen Raum beschränkt bleibt:⁸ sie schädigen Verwandte, Doktoreltern, die ausbildende Universität, Kommilitonen und binden gegebenenfalls Ressourcen, die anderweitig besser investiert worden wären, in Form von Betreuungszeit, Stipendien und anderen Zuwendungen.

All dies können Sie so oder ähnlich in vielen thematisch einschlägigen Publikationen wieder finden. Ich möchte Sie aber nicht ohne einen eigenen originären Gedanken in die Mittagspause entlassen – Plagiate sind noch in einer weiteren Hinsicht schädlich: Sie sind ein Klimakiller, denkt man an Rechnerlaufzeiten, die Publikation und nicht zuletzt den technischen Aufwand, der zunehmend für ihre Aufdeckung betrieben wird.

Die Aufdeckung von Wissenschaftsplagiaten ist im Sinne der akademischen Selbstreinigung ein notwendiges und oft ungewürdigtes – es ist ein verdienstvolles Unterfangen. Eine nachhaltige Strategie zur Förderung guter wissenschaftlicher Praxis ist der notwendigerweise rückwärtsgewandte Blick des Plagiatssuchenden indes nur bedingt. Die Rettung der akademischen Welt bedarf nicht nur einer bewußten Auswahl technischer Mittel, sie bedarf als Gesamtaufgabe einer gelebten Praxis akademischer Integrität.

Mit dem Hinweis auf die gute wissenschaftliche Praxis als *consensus communis* schließt sich nun auch der letzte offene Bogen zu meiner sinnentstellenden Anleihe aus dem *Kommunistischen Manifest* eingangs dieses Vortrages.

FOLIE 20

Ich danke Ihnen für Ihr Interesse und freue mich auf Ihre Fragen.

⁸Allerdings nicht immer, wie dieses tragische Beispiel zeigt: Leung, Pamela T. M., Erin M. Macdonald, Matthew B. Stanbrook, Irfan A. Dhalla, und David N. Juurlink. 2017. A 1980 Letter on the Risk of Opioid Addiction. *The New England journal of medicine* 376 (22): 2194–2195. doi: 10.1056/NEJMc1700150.