Erster Teil: Text- und Data-Mining zwischen Technologie, Innovation und Wissenschaft

Das TDM dient einerseits als eigenständiges Analyseverfahren, das genutzt werden kann um Erkenntnisse über potenziell große Datensätze zu erlangen. Andererseits stellt es die Schlüsseltechnologie für das maschinelle Lernen und andere Teilbereiche des Forschungsfeldes der Künstlichen Intelligenz dar. Im nachfolgenden § 3 werden deswegen sowohl begriffliche als auch technische Grundlagen von TDM und der Künstlichen Intelligenz, eingeschlossen derjenigen Technologien, die für das TDM besonders relevant sind, erläutert. In Bezug auf das Begriffsverständnis des TDM wird zuerst aus technischer und dann aus juristischer Perspektive untersucht, welche Teile des umfangreichen technischen Prozesses unter das TDM zu fassen sind bzw. was *„das"* TDM als solches ist. Daraus resultierend wird ein Begriffsverständnis für die Arbeit gebildet. § 3 stellt daneben auch einen exemplarischen Forschungsprozess vor. Anschließend werden die Muster, die das TDM hervorbringt, mittels einer wissenschaftstheoretischen Herangehensweise kritisch beleuchtet.

Sodann thematisiert § 4 das Verhältnis von Wissenschaft und Urheberrecht, das für die Forschungsmethode des TDMs relevant ist. In diesem Bereich wirken sich bestimmte urheberrechtliche Grundwertungen auf eine Art und Weise aus, die sich vom sonstigen Urheberrecht unterscheidet und die dazu genutzt werden können, wissenschaftliche Forschung zu steuern.

Gerade aufgrund der Relevanz für das Training von KI-Modellen hat das TDM auch eine ökonomische Bedeutung, die es nach Bestrebungen auf europäischer und nationaler Ebene zu nutzen gilt. Aus diesem Grund legt § 5 der Arbeit in § 5 neben dem Verständnis der Steuerungswirkung bzgl. wissenschaftlicher Forschung auch ein Verständnis der Steuerungswirkung des Urheberrechts bzgl. Innovation zugrunde.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

§ 3 Text- und Data-Mining zwischen Künstlicher Intelligenz und digitaler Forschung

A. Das Begriffsverständnis des Text- und Data-Minings

I. Das technische Begriffsverständnis

Unter dem Data-Mining und dem Text-Mining wurde ursprünglich nur die automatisierte algorithmusbasierte Analyse von Daten- oder Textbeständen verstanden, d. h. der eigentliche automatisierte Auswertungsvorgang, nicht hingegen mehrstufige Abläufe mit umfangreichen Vorverarbeitungsschritten wie der Erstellung eines Datensatzes, der als Korpus bezeichnet wird, die vor der eigentlichen Datenanalyse stattfinden.[17] Für die Analyse sind auch die englischen Bezeichnungen der „*knowledge extraction*" (Wissensextrahierung), der „*information discovery*" (Informationsentdeckung) bzw. des „*information harvestings*" und des „*data pattern processings*" (Daten-Muster-Verarbeitung) gebräuchlich.[18] Mehrstufige Verfahren kamen seit den 1990er-Jahren als sog. „*Knowledge Discovery in Databases*" bzw. als Wissensentdeckung in Datenbanken (abgekürzt als KDD) ins Gespräch. Hintergrund war, dass eine zielgerichtete Datenvorbereitung entscheidende Einflüsse auf Analyseergebnisse hat, sie sich aber als äußerst arbeitsintensiv darstellt.[19] Mit dem KDD wurde erstmals ausdrücklich das Ziel verfolgt, Wissen in mehrstufigen Verfahren aus größeren Datengrundlagen zu extrahieren.[20] *Fayyad*, *Piatetsky-Shapiro* und *Smyth*, die das Begriffsverständnis des TDMs prägten, stellten die folgende Definition vor: „*Wissensentdeckung in Datenbanken ist der nicht triviale Prozess der Identifikation gültiger, neuer, potenziell nützlicher und schlussendlich verständlicher*

17 „*In our view, KDD refers to the overall process of discovering useful knowledge from data, and data mining refers to a particular step in this process. Data mining is the application of specific algorithms for extracting patterns from data*.", *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 37, 39, in Bezug auf Text-Mining „*Text mining can be broadly defined as a knowledge-intensive process in which a user interacts with a document collection over time by using a suite of analysis tools*", *Feldman/Sanger,* The Text Mining Handbook, S. 1; ebenso *Beierle/Kern-Isberner,* Methoden wissensbasierter Systeme, S. 145; *Jo,* Text Mining, S. 3.

18 *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 39; *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 178.

19 *Beierle/Kern-Isberner,* Methoden wissensbasierter Systeme, S. 145.

20 *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 39.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Muster in (großen) Datenbeständen".[21] Das KDD wird als ein mehrstufiges Verfahren beschrieben, das es beinhaltet, die Daten zielgerichtet auszuwählen, sie zu vereinheitlichen und zu bereinigen, eine Analysemethode zielgerichtet auszuwählen, sie der eigentlichen algorithmusbasierten Analyse zu unterziehen sowie die Analyseergebnisse im Anschluss auszuwerten.[22] In Abgrenzung dazu solle das Data Mining die algorithmusbasierte Mustererkennung ohne die Vorverarbeitungsschritte bezeichnen.[23] Heute wird das TDM in der Durchführungsweise, die Gegenstand dieser Untersuchung ist, hingegen synonym zum KDD verstanden,[24] wobei die TDM-Analyse nach wie vor den Kern des Prozesses ausmachen soll.[25]

Weiter wird vom sog. „*Information Retrieval*" gesprochen, wenn es darum geht, dass Informationen aus Datenbeständen extrahiert werden sollen. Das Information Retrieval und das Data-Mining werden teilweise dadurch unterschieden, dass das Data-Mining der Auffindung unbekannter Strukturen dient, während das Information Retrieval gerade die Informationen auffinden soll, die bereits bestehen.[26] Ob diese Abgrenzung geeignet ist, ist in Zweifel zu ziehen, schließlich sind, wie später erläutert wird, alle Muster, die aus den Datenbeständen extrahiert werden, in diesen Datenbeständen bereits angelegt.

Weitere Unterfälle sind das sog. „*Web-Mining*" bzw. das „*Web-Crawling*" bzw. „*-Scraping*" oder „*-Harvesting*", bei denen das Web von sog. „*Web-Crawlern*", d. h. speziellen Computerprogrammen, durchsucht wird, die wiederum Daten auslesen und sie „*kratzen*", d. h. herunterladen.[27] Aus technischer Perspektive rufen die Web-Crawler je nach Auftrag eine Viel-

21 *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 40 f.; *Fayyad/Piatetsky-Shapiro/Smyth* in: Fayyad (Hrsg.), Advances in knowledge discovery and data mining, S. 1, 6; übersetzt aus dem Englischen von *Wrobel/Joachims/Morik* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 405, 409.

22 *Beierle/Kern-Isberner,* Methoden wissensbasierter Systeme, S. 147 f.; *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 40 f.

23 *Fayyad/Piatetsky-Shapiro/Smyth* in: Fayyad (Hrsg.), Advances in knowledge discovery and data mining, S. 1, 4.

24 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 177 f.; *Wrobel/Joachims/Morik* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 405, 405, 409.

25 *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 37.

26 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 268.

27 *Jo,* Text Mining, S. 13; *v. Schönfeld,* Screen Scraping und Informationsfreiheit, S. 24 f.; *Elteste*, CR 2015, S. 447, 447; *Schrenk,* Webbots, spiders, and screen scrapers, S. 227.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

zahl an Webseiten auf, die eigentlich für menschliche Nutzer[28] konstruiert wurden,[29] und verarbeiten deren Inhalte.[30] Anschließend laden sie bestimmte Teile dieser Web-Präsenzen herunter, extrahieren in einem weiteren Schritt darin enthaltene Daten und speichern sie dann in einer Meta-Datenbank ab.[31] Web-Crawler und Web-Scraper werden dergestalt unterschieden, dass Erstere auf statische Webseiten zugreifen, deren Inhalt sich nach Schließen der Webseite nicht ändert, wovon Letztere nicht abhängig sind, denn diese vermögen es, Formulare abzusenden und damit menschliches Verhalten zu imitieren.[32] Das Web-Crawling oder -Scraping kann sowohl kommerziellen als auch wissenschaftlichen Zwecken dienen, z. B. der Auswertung von Daten aus sozialen Netzwerken.[33] Grundsätzlich basiert jede Suchmaschine auf Web-Crawlern.[34]

Während *Fayyad*, *Piatetsky-Shapiro* und *Smyth* also davon ausgehen, dass das Data-Mining im engeren Sinne nur die algorithmusbasierte Datenanalyse ist und der Begriff des Wissensentdeckung in Datenbanken wiederum den Gesamtprozess meint, gehen andere davon aus, dass zum Data-Mining nur die Datenbeschaffung zählt.[35] Begrifflich treffen beide Verständnisse zu: Nach Daten zu schürfen hat eine andere Bedeutung als die Daten zu schürfen, also aus ihnen etwas zu generieren. Das gilt auch für das

---

28 Die vorliegende Arbeit bemüht sich um eine geschlechtsneutrale Formulierung. Wo dies nicht gelingt, sind mit der genutzten Form nicht Personen eines bestimmten Geschlechts angesprochen.

29 *Golla/v. Schönfeld*, K&R 2019, S. 15, 16; *H. Ernst/Jochen Schmidt/Beneken,* Grundkurs Informatik, S. 760 f.; *Gumm/Sommer et al*., Einführung in die Informatik, S. 664 f. und deren Erläuterung des technischen Aufbaus von HTML-Webseiten.

30 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 15 ff.

31 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 17 f.; aus technischer Perspektive anschaulich *Schrenk,* Webbots, spiders, and screen scrapers, S. 174.

32 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 17 ff. Damit sind sie technisch anspruchsvoller aufgebaut. Die besondere technische Herausforderung liegt insofern erstens darin, dass Webseiten für menschliche Nutzer programmiert sind, *Golla/v. Schönfeld*, K&R 2019, S. 15, 15, und andererseits, dass insbesondere die Analyse dynamischer Webseiten besondere Herausforderungen birgt.

33 Ein bekanntes Beispiel, in dem Web Scraping zum Einsatz kam, ist die Cambridge Analytica Affäre, in der über die vormals noch funktionsfähige Programmier-Schnittstelle von Facebook Nutzerdaten zu Wahlkampfzwecken ausgelesen wurden, *Golla/v. Schönfeld*, K&R 2019, S. 15, 15; *Kastrenakes,* Facebook will limit developers' access to account data (21.03.2018).

34 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 16 f.; *Schrenk,* Webbots, spiders, and screen scrapers, S. 173; *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 365. Der bekannteste Web-Crawler dürfte damit derjenige sein, den Google für seine Suchmaschine nutzt, vgl. *Golla/v. Schönfeld*, K&R 2019, S. 15, 16.

35 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 177.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Web-Mining, -Crawling oder -Scraping, die zwar der Datenbeschaffung dienen können, bei denen aber bereits die Auswahl dieser Daten auf einer TDM-Analyse basiert.[36]

Das Text-Mining und das Data-Mining unterscheiden sich wiederum hauptsächlich in ihrem Untersuchungsobjekt. Diesbezüglich wird teilweise differenziert, weil Texte als unstrukturiert, während Daten nach mancher Auffassung als strukturiert gelten.[37] Doch auch Texte sind Daten im weiteren Sinne[38] und aufgrund ihrer logischen Binnenstruktur (der Aufbau von Texten, die Gliederung unter Überschriften und in Absätze sowie die Grammatik) nicht in diesem Sinne unstrukturiert.[39] Zudem ist gerade dann, wenn umfassende Vorverarbeitungsschritte durchgeführt werden, kaum eine sinnvolle Abgrenzung mehr möglich, denn während Daten in den Vorverarbeitungsschritten strukturiert werden, wird die Binnenstruktur von Texten teilweise aufgebrochen, wie an späterer Stelle ausführlich erläutert wird.[40]

II. Das urheberrechtliche Begriffsverständnis

Im Urheberrecht hat sich der Begriff des Text- und Data-Minings als ein Sammelbegriff etabliert, der in den jeweiligen Gesetzesmaterialien umschrieben und teils auch legaldefiniert wird.

Die erste deutsche Norm, die das Text- und Data-Mining an urheberrechtlichen Schutzgegenständen ausdrücklich erlaubt, § 60d UrhG a. F., der zwischen dem 01.03.2021 und dem 06.06.2021 galt, verzichtet zwar auf eine Legaldefinition, bezieht sich aber ausweislich der Gesetzesbegründung auf *„einen mehrstufigen Prozess, bei dem große Datenmengen, legaldefiniert als Ursprungsmaterial, in digitaler Form gesammelt, aufbereitet und automatisiert nach bestimmten Merkmalen durchsucht und ausgewertet werden“*.[41] Das Korpus, unter dem § 60d UrhG a. F. die Sammlung der zuvor ver-

36 Davon gehen auch *Golla/v. Schönfeld*, K&R 2019, S. 15, 17 f. aus.
37 *Käde*, Kreative Maschinen und Urheberrecht, S. 57; *W. Zimmer*, Ansturm der Algorithmen, S. 196; anderer Auffassung ist *Jo*, Text Mining, S. 3, demzufolge das Text-Mining ein Spezialfall des Data-Mining ist.
38 Dazu näher S. 38.
39 *Heyer/Quasthoff/Wittig*, Text Mining: Wissensrohstoff Text, S. 1; *Jo*, Text Mining, S. 3, *W. Zimmer*, Ansturm der Algorithmen, S. 196 f.; Gesellschaft für Informatik, Text Mining.
40 Vgl. dazu S. 48 ff.
41 BT-Drs. 18/12329, S. 40.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

vielfältigten, normalisierten, strukturierten und kategorisierten Inhalte versteht, wird nach dem gesetzgeberischen Verständnis anschließend – u. U. mit eigens programmierten Computerprogrammen – ausgewertet.[42] Diese extrahieren *„statistische Häufigkeiten oder Korrelationen"*, die schließlich für die wissenschaftliche Forschung zur Verfügung stehen.[43] Ziel des Gesamtprozesses sei es, *„neue Erkenntnisse zu gewinnen"*.[44] Dieses weit gefasste Verständnis lässt zunächst die Schlussfolgerung zu, der Gesetzgeber meine mit dem TDM den Gesamtprozess, d. h. die Datenaufbereitung und die -Analyse und vertrete insoweit ein an das KDD angelehntes Verständnis. Ein anderer Eindruck wird vermittelt, weil die Analyse-Software als *„eigentliche Software für das sogenannte Text und Data Mining"* bezeichnet wird,[45] das wiederum die Schlussfolgerung zulässt, das TDM sei nur die algorithmusbasierte Korpusanalyse exklusive der Vorverarbeitungsschritte. An dieser Uneindeutigkeit zeigt sich der vorläufige Charakter der Norm, die in dem Bewusstsein, bald neues Unionsrecht umsetzen zu müssen, verabschiedet wurde.[46]

Auch der DSM-RL vom 17.04.2019 liegt ein weites Verständnis des TDMs zugrunde, indem sie es in Art. 2 Nr. 2 DSM-RL als *„eine Technik für die automatisierte Analyse von Texten und Daten in digitaler Form, mit deren Hilfe Informationen unter anderem — aber nicht ausschließlich — über Muster, Trends und Korrelationen gewonnen werden können"* legaldefiniert. Demnach dient das TDM dem Zweck, digitale Informationen computergestützt auszuwerten, um neue Erkenntnisse zu gewinnen und Trends zu erkennen.[47] Doch ist daraus nicht zwingend zu schlussfolgern, dass die DSM-RL mit dem Begriff des TDMs den Gesamtprozess meint, so erlauben Art. 3 und 4 DSM-RL Vervielfältigungen (und Entnahmen) für das TDM.[48] Das lässt den Schluss zu, dass das TDM die eigentliche Analyse ist,

42 BT-Drs. 18/12329, S. 40.
43 BT-Drs. 18/12329, S. 40.
44 BT-Drs. 18/12329, S. 22.
45 BT-Drs. 18/12329, S. 40.
46 BT-Drs. 18/12329, S. 40.
47 DSM-RL Erw. 8 S. 1, 2.
48 In der deutschen Übersetzung enthält Art. 3 Abs. 1 DSM-RL einen grammatikalischen Fehler: *„Die Mitgliedstaaten sehen eine Ausnahme […] für Vervielfältigungen und Entnahmen vor, die […] zum Zwecke der wissenschaftlichen Forschung für die Text und Data Mining vorgenommen werden"*. In Art. 4 Abs. 1 DSM-RL wird es demgegenüber klarer, dieser sieht Ausnahmen oder Beschränkungen *„für zum Zwecke des Text und Data Mining"* vorgenommenen Vervielfältigungen bzw. Entnahmen vor. In der englischen Sprachfassung des Art. 3 Abs. 1 DSM-RL heißt es: *„in order to carry out, for the purposes of scientific research, text and data mining"*.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

für die im Vorfeld bestimmte Handlungen – die Vorverarbeitungsschritte – erforderlich sind.

Die aktuellste Normfassung im deutschen Urheberrecht kombiniert eine reformierte Fassung des § 60d UrhG mit dem neu geschaffenen § 44b UrhG, der das TDM in Anlehnung an das Verständnis von Art. 2 Nr. 2 DSM-RL als *„die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen"* definiert. *„Für das Text und Data Mining"*[49] dürfen wiederum bestimmte urheberrechtlich relevante Handlungen vorgenommen werden, §§ 44b Abs. 2, 60d Abs. 1 UrhG. Auch diesbezüglich liegt es also nahe, dass der Gesetzgeber unter dem TDM nur die algorithmusbasierte Analyse versteht und verschiedene Handlungen zu diesem Zwecke, d. h. die Vorverarbeitungsschritte, erlauben möchte.

Neben diesen Abgrenzungen deuten die gesetzgeberischen Formulierungen und die Ausführungen der DSM-RL darauf hin, dass sich die Normgeber auf das TDM als Forschungsmethode mit einem konkreten Analyseziel beziehen, wie sie heute z. B. in den digitalen Geisteswissenschaften (englisch: Digital Humanities) praktiziert wird.[50] Dennoch ist der Begriff so weit gefasst, dass auch eine große Bandbreite anderer Anwendungen, die auf der Technologie des TDM basieren, erfasst werden, wie an späterer Stelle diskutiert wird.[51]

III. Begriffsfindung

Aus technischer und rechtlicher Perspektive zeigt sich, dass mit dem Begriff des TDMs sowohl die eigentliche algorithmusbasierte Datenanalyse als auch der Gesamtprozess des TDMs inklusive der Datenvorverarbeitung adressiert sein kann. Gerade aus technischer Perspektive existiert eine Fülle wechselbezüglicher Begriffe mit weitreichenden inhaltlichen Überschneidungen. Die verschiedenen Definitionen haben im Wesentlichen gemeinsam, dass sie sich darauf beziehen, dass aus mehr oder minder strukturier-

49 Im Gesetz wird statt der in dieser Arbeit verwendeten Schreibweise – Text- und Data-Mining – eine Schreibweise ohne Bindestriche vorgenommen, das gilt sowohl im deutschen Recht als auch in der DSM-RL. Wenn aus dem Gesetz zitiert wird, werden deswegen keine Bindestriche genutzt.

50 Das wird ausführlich auf S. 36 ff. thematisiert.

51 Vgl. dazu den zweiten Teil der Arbeit, insbesondere § 9.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

ten Daten ein Mehrwert gezogen wird, der in der Offenbarung neuer Erkenntnisse, unbekannter Muster oder Korrelationen liegen kann. Uneinigkeiten bestehen darin, welche Arbeitsschritte für das TDM begriffsprägend sind. Der Begriff des TDMs ist im Urheberrecht ebenso uneindeutig, aber doch tendenziell weit gefasst.

Trotz der Uneindeutigkeiten lässt sich eine gemeinsame Linie der juristischen und technischen Perspektive herausarbeiten, die mehrstufige Verfahren unter den Gesamtkomplex des TDMs fassen, bei dem große Datenbestände aufgebaut und aufbereitet werden, um sie anschließend auf Muster oder Strukturen hin zu analysieren.

Der folgenden Untersuchung wird ein Verständnis zugrunde liegen, das das TDM als algorithmusbasierte Methode zur Mustererkennung ansieht, das zu verschiedensten Zwecken genutzt werden kann und das eine umfassende Datenvorbereitung erfordern kann. Das TDM erfasst dabei in Anlehnung an den Begriff des KDDs nicht nur die eigentliche Datenanalyse, sondern den Gesamtprozess, wobei bei Bedarf zwischen den Vorverarbeitungsschritten und der algorithmusbasierten TDM-Analyse im eigentlichen Sinne differenziert wird. Wenn diese Analyse gemeint ist, wird dies durch die Formulierung „*TDM im engeren Sinne*“ oder „*TDM-Analyse im engeren Sinne*“ gesondert klargestellt. Das Text-Mining wird in dieser Untersuchung als Unterkategorie des Data-Minings angesehen, wobei die Begriffe nur in den Fällen getrennt verwendet werden sollten, wenn es auf ihre Differenzierung ankommt.

B. Text- und Data-Mining zur Erkenntnisgewinnung

Seine Anfänge fand das TDM als geisteswissenschaftliches Forschungswerkzeug bereits in den 1940er-Jahren, als der italienische Jesuit *Roberto Busa*, der noch im Rahmen seiner Dissertation zur Beantwortung seiner Forschungsfragen händisch Wörter zählte, in Kooperation mit dem heutigen Technologie- und Softwarekonzern, damals aber noch Lochkarten produzierenden, Unternehmen IBM, zur Erforschung unbewusster Formulierungen das Gesamtwerk von *Thomas von Aquin* systematisierte.[52] Das Projekt sollte es ermöglichen, in umfangreichen Textsammlungen schnell bestimmte Phänomene oder Passagen aufzufinden.[53] Aus diesem Projekt

52 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 3.

53 *Sasaki* in: Klimpel/Euler (Hrsg.), Der Vergangenheit eine Zukunft, S. 84, 85. Ähnliche Ansätze gab es auch in den Vereinigten Staaten, diese konnten aber erst in

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

ist der *Index Thomisticus*,[54] ein noch heute erforschtes Korpus hervorgegangen, bestehend aus dem systematisierten, lemmatisierten und annotierten Gesamtwerk von *Thomas von Aquin* sowie aus Sekundärquellen. *Roberto Busa* musste also feststellen, dass technische Hilfe notwendig ist, um überhaupt bzw. in einem angemessenen Zeitrahmen und Budget die nicht offensichtlichen Zusammenhänge und Bedeutungen in größeren Textbeständen zu erkennen.

In Bezug auf die digitalen Geisteswissenschaften ist viel diskutiert, ob es sich um eine Querschnittsmaterie handelt, die computergestützt neue Erkenntnisse in geisteswissenschaftlichen Fragestellungen untersucht, oder um eine eigene Forschungsdisziplin.[55] Letzteres würde der Vielfalt der geisteswissenschaftlichen Disziplinen nicht gerecht.[56] Allerdings sind auch die Fragestellungen der digitalen Geisteswissenschaften im Vergleich zu denjenigen der traditionellen Geisteswissenschaften durch die digitalen

---

den 1960er Jahren realisiert werden, *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 3.

54 Dieser ist heute auch online verfügbar, s. dazu https://www.corpusthomisticum.org/wintrode.html.

55 *Sahle* in: Baum/Stäcker (Hrsg.), Grenzen und Möglichkeiten der Digital Humanities: „*Die DH erscheinen als voll ausgereiftes, eigenständiges Fach. Es kann kaum bezweifelt werden, dass wir es hier mit einer eigenständigen Disziplin zu tun haben. Allerdings scheint die Sache doch nicht ganz so einfach zu sein, weil diese Disziplin einen besonderen Charakter hat. Als »Brückenfach«, ihrem Wesen nach inter- oder transdisziplinär und zugleich viele spezialisiertere Ausprägungen umgreifend, bestimmt sie sich nämlich nicht nur in Abgrenzung von anderen Fächern, sondern gerade auch in ihrer Hinwendung zu ihnen.*". Anderer Auffassung sind *Jannidis/Kohle/Rehbein*, die die digitalen Geisteswissenschaften als die Summe aller Versuche definieren, „*die Informationstechniken auf den Gegenstandsbereich der Geisteswissenschaften zu übertragen*", *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 13, die aber auch die Eigenheiten der untersuchten Fragestellungen als kennzeichnend für die Anerkennung einer eigenen Forschungsdisziplin ansehen, *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 13 f. Das soll gerade dann gelten, wenn Erkenntnisse getroffen werden sollen, die ohne die Mittel der digitalen Geisteswissenschaften nicht auffindbar wären. Gerade das Faktum, dass es nicht „*die analogen Geisteswissenschaften*" gibt, spricht auf der anderen Seite dafür, auch die digitalen Geisteswissenschaften nicht als Forschungsdisziplin einzuordnen, *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 15. Diese „*Werkzeugtheorie*" spricht gleichwohl nicht gegen die Bedeutsamkeit oder das Potenzial der digitalen Geisteswissenschaften.

56 Das repräsentiert auch das Tagungsprogramm der Jahrestagungen des Verbands Digital Humanities im deutschsprachigen Raum e.V. (DHd), die die wohl größten jährlichen Tagungen in diesem Themenfeld im deutschsprachigen Raum abhält sowie das Selbstverständnis dieses Verbands, Digital Humanities im deutschsprachigen Raum, Was sind digitale Geisteswissenschaften?

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Mittel bestimmt, weswegen die digitalen Geisteswissenschaften auch nicht auf eine reine Methodik reduziert werden sollten.[57] Letztlich fällt es schwer, die digitalen Geisteswissenschaften begrifflich zu umreißen. Jedenfalls wird das TDM in den digitalen Geisteswissenschaften als Schlüsseltechnologie genutzt, um auch – aber nicht nur – sprachbasierte Inhalte auszuwerten.[58]

C. Text- und Data-Mining und Künstliche Intelligenz

Wie bereits erläutert wurde, ist das TDM eng mit dem Forschungsfeld der KI verbunden. Im Folgenden werden zunächst deren Begriffsverständnis und anschließend verschiedene Ausprägungen bzw. Teilbereiche, die besonders relevant für das TDM sind, thematisiert, bevor das Verhältnis von TDM und KI erläutert werden kann.

I. Technische Grundlagen der Künstlichen Intelligenz

1. Begriffsentwicklung der Künstlichen Intelligenz

Der Begriff der Künstlichen Intelligenz stammt aus dem Jahr 1955 und wurde durch *McCarthy* und *Minsky* geprägt.[59] Die theoretischen Grundsteine für einzelne Technologien wurden aber bereits in den 1940er-Jahren gelegt.[60] Trotz dieser frühen Anfänge wurden besondere Fortschritte erst deutlich später aufgrund der fortschreitenden Rechenleistung von Großcomputern zu verzeichnet, was auch die Nutzung künstlicher neuronaler Netze bzw. des Deep Learnings ermöglichte.[61]

Der Begriff der KI sowie die Beantwortung der Frage, wann ein Programm als intelligent bezeichnet werden kann, ist bis heute äußerst

57 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 13 f.
58 *„Zentraler Arbeitsbereich"*, *Döhl*, RuZ 2020, S. 195, 198.
59 *McCarthy/Minsky/Rochester/Shannon*, AI Mag 2006, S. 12 ff.; *Pallay* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 13; *Kaplan*, Künstliche Intelligenz, S. 27 f.; *Lenzen*, Künstliche Intelligenz, S. 22.
60 *Pallay* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 5 f.; *Kaplan*, Künstliche Intelligenz, S. 27 ff.; *Negnevitsky*, Artificial intelligence, S. 2.
61 *Mikut*, at 2020, S. 295, 295 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

strittig.[62] Bzgl. der Bestimmung des Autonomiegrads von Systemen, die auf Technologien der KI basieren, wird an die Differenzierung zwischen der sog. „*starken*" und „*schwachen*" KI angeknüpft, die wiederum davon abhängt, ob die entsprechenden Modelle universell einsetzbar sind und ein Bewusstsein haben oder lediglich die Intelligenz des Menschen imitieren.[63] Es sei an dieser Stelle klargestellt, dass es eine „*starke* – also eine universell einsetzbare – KI nach dem aktuellen Stand der Technik nicht gibt.[64] Die begrifflichen Schwierigkeiten zeigen sich auch in Definitionen oder Definitionsversuchen der deutschen und europäischen Normgeber oder Exekutivorgane. Die Europäische Kommission versteht unter der Künstlichen Intelligenz Systeme, die auf verschiedenen Analysen basierend

62 „*Der Begriff der künstlichen Intelligenz ist dabei ebenso schillernd wie konturenlos*", *Specht*, GRUR 2019, S. 253, 253. Eine Sammlung an 70 Definitionen der Intelligenz bieten *Legg/Hutter* in: Goertzel/Wang (Hrsg.), Advances in artificial general intelligence, S. 17 ff., 2.

63 *Hauck/Cevc*, ZGE 2019, S. 135, 140 f. präzisieren den Begriff der schwachen KI hin zu einer spezifischen bzw. einer anwendungsbereichsbezogenenen KI; zum Kriterium der Autonomie ausführlich *Kirn/Müller-Hengstenberg*, MMR 2014, S. 225, 225; *Pieper*, InTeR 2018, S. 9 ff., 11; *Aust,* Das Zeitalter der Daten, S. 26 f.; *Grätz,* Künstliche Intelligenz im Urheberrecht, S. 4 Fn. 31, 13 f.; *Kaplan,* Künstliche Intelligenz, S. 83; *Käde,* Kreative Maschinen und Urheberrecht, S. 41; *Mainzer,* Künstliches Leben und virtuelle Agenten (16.01.1997); in Bezug auf denkende Maschinen *Strube/Ferstl/Konieczny/Ragni* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 21, 44. Auf letzteren Standpunkt stellt sich die deutsche Bundesregierung, Bundesregierung, Strategie Künstliche Intelligenz der Bundesregierung (2018), S. 4 f.

64 *Hauck/Cevc*, ZGE 2019, S. 135, 140 f.; *Reichwald/Pfisterer*, CR 2016, S. 208, 210 f.; *Niederée/Neidl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 2; *Bauberger/B. Beck/Burchardt/Remmers* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 907, 507; *Lenzen,* Künstliche Intelligenz: Fakten, Chancen, Risiken, S. 199 f.; *Strube/Ferstl/Konieczny/Ragni* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 21, 44 betonen, dass bisher keine autonom denkenden Maschinen entwickelt werden konnten; ebenso *Volland,* Die kreative Macht der Maschinen, S. 14. Optimistisch äußert sich *Pieper*, InTeR 2018, S. 9 ff., 14: „*Es ist anzunehmen, dass es in Zukunft autonome Systeme geben wird, die mit „Starker KI" ausgestattet sind*", zweifelnder ist *Grätz,* Künstliche Intelligenz im Urheberrecht, S. 14: „*Ungewiss ist, ob dies jemals gelingen wird*". Ein Forschungszweig der KI richtet sich nach der Suche dieser „*künstlichen allgemeinen Intelligenz*", *Lenzen,* Künstliche Intelligenz, S. 32 f., 120: „*Zu prognostizieren, wie lange es noch dauern wird, eine solche künstliche allgemeine Intelligenz zu realisieren, ist zu einem beliebten Gesellschaftsspiel geworden. Die Antworten reichen von »unmöglich« bis »unvermeidlich«, der Zeithorizont wird mal mit 20, mal mit 200 Jahren angegeben.*"

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

(teil-)autonome, intelligente Verhaltensweisen zeigen.[65] Angeknüpft wird also an bestimmte Eigenschaften. Der KI-Verordnungsentwurf zur „*Festlegung harmonisierter Vorschriften für Künstliche Intelligenz*" definiert ein System „*der Künstlichen Intelligenz*" in ihrem Art. 3 Nr. 1 als „*Software, die mit einer oder mehreren der im Anhang I aufgeführten Techniken und Konzepte entwickelt worden ist und im Hinblick auf eine Reihe von Zielen, die vom Menschen festgelegt werden, Ergebnisse wie Inhalte, Vorhersagen, Empfehlungen oder Entscheidungen hervorbringen kann, die das Umfeld beeinflussen, mit dem sie interagieren*". Der Anhang dieses Verordnungsentwurfs benennt das maschinelle Lernen inklusive des Deep Learnings sowie logik- und wissensgestützte Systeme und statistische Ansätze. Auch die deutsche Bundesregierung wählt in ihrer KI-Strategie ein beispielbasiertes Verständnis und versteht jedenfalls unter schwacher KI – „*Deduktionssysteme […], wissensbasierte Systeme […], Musteranalysen und Mustererkennung […], Robotik*" und „*intelligente multimodale Mensch-Maschine-Interaktionen*".[66] Mittels dieser technologiespezifischen Aufzählungen wird es vermieden, definitorisch zu bestimmen, was die KI eigentlich ausmacht.

Die Schwierigkeiten, die KI abstrakt zu definieren, sind mitunter darauf zurückzuführen, dass es kein allgemeingültiges Verständnis von „*Natürlicher Intelligenz*" oder von „*Intelligenz*" als solcher gibt.[67] Die Antwort, die der Duden auf die Frage nach dem Begriff der „*Intelligenz*" liefert, ist für diese Kernproblematik bezeichnend: Er versteht darunter „*die Fähigkeit [des Menschen], abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten*".[68] Das wirft die Frage auf, ob Intelligenz zwangsläufig mit dem Menschsein zusammenhängt. Das war auch das ursprüngliche Verständnis von Künstlicher Intelligenz, es galt also als intelligent, was sich zumindest als Mensch ausgeben konnte.[69] Jedenfalls die Lernfähigkeit

---

65 Europäische Kommission, Künstliche Intelligenz für Europa - COM(2018) 237 final, S. 1. Eine gute Übersicht zu den verschiedenen europäischen Bestrebungen für den Datenmarkt und die Bereiche der KI bietet *Jüngling*, MMR 2020, S. 440 ff.

66 Bundesregierung, Strategie Künstliche Intelligenz der Bundesregierung (November 2018), S. 4 f.

67 *Kaplan*, Künstliche Intelligenz, S. 15.

68 Duden, Definition der Intelligenz, abrufbar unter: https://www.duden.de/rechtschreibung/Intelligenz.

69 *McCarthy/Minsky/Rochester/Shannon*, AI Mag 2006, S. 12, 12; *Turing* in: Carpenter/Doran, A. M. Turing´s ACE report of 1946 and other papers, S. 106 (123); *Pallay* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 9; *Turing* in: Carpenter/Doran (Hrsg.), A. M. Turing´s ACE report of 1946 and other papers, S. 106 ff., 123; *Kaplan*, Künstliche Intelligenz, S. 15. In diesem Kontext sei auch der nach *Alan Turing* benannte *Turing Test* erwähnt, der in sei-

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

kann nicht allein als Kriterium für Intelligenz fungieren und auch auf auf bloße mathematische Leistungsfähigkeit kann es nicht ankommen, andernfalls basierte bereits jeder Taschenrechner auf Künstlicher Intelligenz.[70] Auch wenn Intelligenz an Anpassungsfähigkeit an neue Situationen oder eine breite Einsatzfähigkeit gekoppelt wäre, wäre KI nach aktuellem Stand nicht als intelligent zu bezeichnen, schließlich gibt es noch keine starke KI.

Anhand dieser Beispiele wird deutlich, wie schwierig die Begriffsfindung fällt. Ein wesentliches Kriterium eines Systems der KI ist jedoch seine Anpassungsfähigkeit bzw. die Lernfähigkeit zur Verbesserung der getroffenen Entscheidungen,[71] d. h. die Rückschlüsse, die aus Datenanalysen gezogen werden, müssen sich jedenfalls theoretisch auf zukünftige Analysen auswirken und auf neue Sachverhalte angewendet werden können.[72]

2. Die Technologie des maschinellen Lernens

Ein wichtiger Teilbereich der KI ist das maschinelle Lernen, das häufig fälschlicherweise synonym zur KI verstanden wird.[73] Beim maschinellen Lernen lernen Algorithmen anhand eines möglichst großen Erfahrungsschatzes, komplexe Probleme zu lösen, indem abstrakte Regeln gebildet werden, die sich aus Mustern ergeben, die in den analysierten Daten enthalten sind – gelernt wird also anhand von Beispielen.[74] Der Lernprozess kann entweder überwacht oder unüberwacht oder in einer Abstufung erfolgen. Sofern das Programm einen Datensatz zum Training erhält, der bereits Kategorisierungen enthält, spricht man von überwachtem Lernen.[75]

---

nem Ursprung ein Gedankenspiel war, das prüfen sollte, ob eine Maschine dazu in der Lage sei, einen Menschen in einem fünfminütigen fiktiven Gespräch über sein Dasein als Maschine täuschen kann. Falls die Maschine dazu in der Lage sein sollte, solle sie als intelligent gelten, *Pallay* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 11 ff.

70 *Kaplan,* Künstliche Intelligenz, S. 16; *Lenzen,* Künstliche Intelligenz, S. 29.

71 *Ertel,* Grundkurs Künstliche Intelligenz, S. 3; *Zech*, ZfPW 2019, S. 198, 200.

72 *Niederée/Neidl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 5 f.

73 *Ertel,* Grundkurs Künstliche Intelligenz, S. 201.

74 *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 21, 24 ff.; „*Computerprogramme lernen, indem sie Muster aus Daten extrahieren*", *Kaplan,* Künstliche Intelligenz, S. 44.

75 *Stiemerling*, CR 2015, S. 762, 763; *Bramer,* Principles of Data Mining, S. 4 f.; *Lenzen,* Künstliche Intelligenz, S. 50 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

In dem Fall lernt das Programm anhand des kategorisierten Trainingsmaterials, bei welchen Kategorisierungen in dem Trainingsmaterial welche Merkmale auftreten.[76] Nach Abschluss der Trainingsphase wird der trainierte Algorithmus auch als „*Modell*" bezeichnet.[77] Anschließend kann das Modell anhand des gewonnenen Kenntnisschatzes an nicht kategorisierten Daten erkennen, welches Wort oder Bild welcher Kategorie zuzuordnen ist. Im Rahmen des unüberwachten Lernens hingegen extrahiert das Programm aus nicht-kategorisierten Daten eigenständige Regeln, Gruppen, Kategorien, Muster oder Zusammenhänge aus den nicht kategorisierten Trainingsdaten.[78] Daneben sei das verstärkende Lernen erwähnt, bei dem die künstlichen Neuronen im Rahmen von eigentlich unüberwachtem Lernen Rückmeldungen erhalten, die ihre Kategorisierung stützen oder ablehnen, wodurch Kategorisierungen angepasst und Entscheidungsprozesse verbessert werden.[79]

In besonderer Weise eignet sich das maschinelle Lernen für die Verarbeitung der natürlichen Sprache, die aufgrund der Vielschichtigkeit der Sprache – Syntax, Semantik und Pragmatik – technisch besonders anspruchsvoll zu erfassen ist.[80] Bei gesprochener Sprache kommen des Weiteren die Phonologie, d. h. das akustische Verständnis bei aussprachebedingten Ungenauigkeiten, sowie das Verständnis der Intonation, der Einordnung bewusster sprachlicher Stilmittel und dem Herausfiltern von Umgebungsgeräuschen hinzu.[81] Die Technik steht daher vor der Herausforderung, den Bedeutungsgehalt abhängig von Kontext, Satzgefüge, Satz-

---

76 *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 40; *Kaplan,* Künstliche Intelligenz, S. 46 f.; *Lenzen,* Künstliche Intelligenz, S. 50 f.

77 *Drexl/Hilty et al*., Technical Aspects of Artificial Intelligence: An Understanding from an Intellectual Property Perspective (2019), S. 5.

78 *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 41; *Lenzen,* Künstliche Intelligenz, S. 51; *Mallot/Hübner* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 357, 380 f.

79 Vgl. dazu *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 14; *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 44; *Lenzen,* Künstliche Intelligenz, S. 51.

80 Andere Algorithmen haben dabei besondere Schwierigkeiten, was u. a. bei Sprachassistenten ersichtlich wird, die erst in den letzten Jahren Fortschritte machten und zuvor nur sehr rudimentär einsetzbar waren. Vgl. zu dieser Problematik *Schaffranietz/F. Neumann* in: Keuper/Neumann (Hrsg.), Wissens- und Informationsmanagement, S. 149, 154; anschaulich erläutert von *W. Zimmer,* Ansturm der Algorithmen, S. 197 f.; *R. T. Kreutzer/Sirrenberg,* Künstliche Intelligenz verstehen, S. 30; *Nilsson,* The quest for artificial intelligence, S. 141 f.

81 *Kaplan,* Künstliche Intelligenz, S. 73.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

zeichen, Empfängerkreis sowie akustischen Besonderheiten sowie Humor, Ironie und Sarkasmus und anderen sprachlichen Mitteln und Trägern subtiler sprachlicher Bedeutungen zu verstehen.[82] Zudem stellt es eine Herausforderung dar, Sprache insbesondere aus handgeschriebenen oder aus schlecht erhaltenen Schriftstücken zu extrahieren.[83] Das gesamte Anwendungsfeld wird unter der natürlichen Sprachverarbeitung (englisch: *Natural Language Processing* (NLP)) eingeordnet und im Forschungsfeld der Computerlinguistik untersucht.

Das für die Verarbeitung natürlicher Sprache sowie für Bild- und Mustererkennung momentan am besten geeignete Verfahren aus dem Bereich des maschinellen Lernens ist das bereits erwähnte Deep Learning,[84] das sich in den letzten Jahren aufgrund der sich laufend verbessernden Rechenkapazitäten stetig weiterentwickelt hat,[85] sodass es heute praktisch anwendbar ist. Ziel war ursprünglich, die Funktionsweise des menschlichen Gehirns zu imitieren.[86]

Nach dem Vorbild der Neuronen des menschlichen Gehirns basieren Deep Learning Algorithmen auf künstliche neuronale Netze, die mittels künstlicher Neuronen in einer Netzstruktur programmiert sind.[87] Jedes der künstlichen Neuronen ist dabei eine mathematische Gleichung, die für die Kategorisierung bzw. Klassifizierung von Merkmalen geschrieben wurde.[88] Unterschieden werden die Eingabeschicht (englisch: *input layer*), eine variable Zahl an verdeckten Schichten (englisch: *hidden layer*) und eine Ausgabeschicht (englisch: *output layer*).[89] Die Funktionsweise eines künstlichen neuronalen Netzes lässt sich am besten dergestalt beschreiben,

---

82 Vgl. dazu auch *R. T. Kreutzer/Sirrenberg,* Künstliche Intelligenz verstehen, S. 29; *Schaffranietz/F. Neumann* in: Keuper/Neumann (Hrsg.), Wissens- und Informationsmanagement, S. 149, 153; *R. T. Kreutzer/Sirrenberg,* Künstliche Intelligenz verstehen, S. 30.

83 Dagegen ist es mit aktueller OCR-Software einfacher, den Text aus Scans computer- oder schreibmaschinengeschriebener Texte zu extrahieren.

84 *Jannidis/Kohle/Rehbein,* Digital Humanities, S. 290; *Ertel,* Grundkurs Künstliche Intelligenz, S. 330.

85 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 176; *Styczynski/Rudion/Naumann,* Einführung in Expertensysteme, S. 135.

86 *Kaplan,* Künstliche Intelligenz, S. 45; *Styczynski/Rudion/Naumann,* Einführung in Expertensysteme, S. 133 f.

87 *Jüngling,* MMR 2020, S. 440, 331; *Ertel,* Grundkurs Künstliche Intelligenz, S. 285 f.; *Kaplan,* Künstliche Intelligenz, S. 45 f.; *H. Ernst/Jochen Schmidt/Beneken,* Grundkurs Informatik, S. 801.

88 *Mallot/Hübner* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 357, 360 ff., 365 ff.

89 *Kaplan,* Künstliche Intelligenz, S. 46.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

dass die einzelnen Neuronen jeweils Analysen, z. B. Kategorisierungen, vornehmen und dem Ergebnis ihrer Analyse einen Wert zuordnen. In der Eingabeschicht werden dabei die Eingabedaten verarbeitet.[90] Dabei wird die Schlussfolgerung aus dem Lernmaterial extrahiert und losgelöst von den Trainingsdaten in den sog. Gewichten abgespeichert.[91] Das Ergebnis sowie das Gewicht, das die Wahrscheinlichkeit ausdrückt, mit der die Entscheidung getroffen wurde, werden anschließend an das nachgeschaltete Neuron weitergegeben.[92] Anschließend wird die zweite Schicht aktiviert, die dann wiederum die dritte Schicht aktiviert. Die weitergegebenen Werte bilden für die nächste Schicht stets die Grundlage, auf der dann eine genauere Entscheidung getroffen wird.[93] In der Ausgabeschicht wird letztlich das Endergebnis ausgegeben.[94]

Eine große Zahl an Schichten bildet dabei ein tiefes neuronales Netzwerk, woher der Begriff des *„deep learnings"*, also des tiefen Lernens, stammt.[95] Die große Anzahl an Schichten sowie Rückkoppelungen und Querverschaltungen der Neuronen verbessern dabei die Genauigkeit,[96] denn anhand einer großen Fülle an Beispielen und Ergebnissen lassen sich allgemeingültige und genauere Regeln erzielen.[97] Während die Lernfähigkeit des Gehirns in der Fähigkeit der Neuronen zur Adaption mittels einer

90 *Söbbing*, MMR 2021, S. 111, 112.
91 *Ertel*, Grundkurs Künstliche Intelligenz, S. 342.
92 *Lenzen*, Künstliche Intelligenz, S. 53.
93 Bei der sog. Backpropagation werden die Gewichte auch nachträglich verändert, und Änderungswünsche in der Kategorisierung zurückgegeben, um künftige Entscheidungen zu verbessern *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 69 f.
94 *Söbbing*, MMR 2021, S. 111, 112.
95 *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 62.
96 Bei sog. Feedforward-Netzen erfolgt die Informationsweitergabe nur in eine Richtung, noch nützlicher sind allerdings die Backpropagation-Netze, bei denen eine Rückleitung der Information erfolgt. Außerdem ist es technisch möglich, die Neuronen, die in derselben Schicht verschalten sind, miteinander zu verknüpfen; *Lenzen*, Künstliche Intelligenz, S. 56. Zudem können sich auch die Verschaltungen bzw. die Verknüpfungen der Neuronen sowie die Aktivierungsfunktionen anpassen, *Styczynski/Rudion/Naumann*, Einführung in Expertensysteme, S. 147 f.
97 *Stiemerling*, CR 2015, S. 762, 764; *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 16.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

wandelbaren Leitfähigkeit der Synapsen liegt,[98] erfolgt das „*Lernen*" der künstlichen Neuronen u. a. in der Anpassung der Gewichte.[99]

3. Wissensbasierte Systeme und Expertensysteme

Einen weiteren Anwendungsbereich der KI bilden sog. Expertensysteme, die als wissensbasierte Systeme, die die Kennzeichen menschlicher Experten bzw. Expertise nachbilden[100] und die als technische Unterstützung in verschiedenen Entscheidungsfindungen genutzt werden.[101] Die Vorläufer der wissensbasierten Systeme fanden ihren Anfang bereits in den 1970er-Jahren,[102] die heutigen Expertensysteme basieren allerdings auf Methoden der KI.[103] Die technische Schwierigkeit ist dabei, das Expertenwissen nachzubilden, das sich normalerweise aus einer Fülle an Erfahrungen und der sich daraus entwickelnden Intuition entwickelt.[104] Technisch wird das erreicht, indem aus der Wissensdatenbank allgemeine Regeln extrahiert werden, die wiederum auf Einzelfälle anwendbar sind und Wahrschein-

98 *Ertel*, Grundkurs Künstliche Intelligenz, S. 288.

99 *Stiemerling*, CR 2015, S. 762, 764; *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 29.

100 *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 11 ff.; *Styczynski/Rudion/Naumann*, Einführung in Expertensysteme, S. 10 f.

101 „*Vernünftigerweise*", *Strube/Ferstl/Konieczny/Ragni* in: Görz/Schneeberger/Schmid (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 21, 22.

102 Die wissensbasierten Methoden stießen allerdings schnell an ihre Grenzen und fielen insofern auch in den sog. „*AI-Winter*", d. h. einer Ernüchterungsphase dahingehend, was von KI-Systemen erwartet werden kann, *Görz/Braun/Schmid* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 1, 8.

103 *Leimeister*, Einführung in die Wirtschaftsinformatik, S. 261.

104 *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 11 ff.; *Görz/Braun/Schmid* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 1 ff., 14 erklären diesbezüglich anschaulich, wodurch eine menschliche Expertise zustande kommt: „*Menschliches Problemlösen zeichnet sich dadurch aus, dass das dabei verwendete Wissen zumeist vage und unvollständig ist. Die Qualität menschlicher Problemlöser zeigt sich gerade darin, dass und wie sie unerwartete Effekte und Ausnahmesituationen aufgrund ihrer Berufserfahrung bewältigen können, dass sie aus Erfahrung lernen, ihr Wissen also ständig erweitern, und dass sie aus allgemeinem Wissen nicht nur nach festen Schlussregeln, sondern auch durch Analogie und mit Intuition Folgerungen gewinnen*"; ähnlich auch *Nebel/Wölfl* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 27, 110 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

lichkeiten sowie Handlungsvorschläge ermitteln,[105] die z. B. für medizinische Diagnosen nutzbar sind.[106]

II. Überschneidungen von Text- und Data-Mining mit Technologien der Künstlichen Intelligenz

Die Frage, wie sich das Forschungsfeld der KI und die Technologie des TDMs zueinander verhalten, wird angesichts der begrifflichen Unschärfen uneinheitlich beantwortet. Eine Schnittmenge im Bereich der KI ergibt sich mit dem Teilbereich des maschinellen Lernens und mit sog. Expertensystemen oder wissensbasierten Systemen. Einigkeit besteht darin, dass jedenfalls das maschinelle Lernen verwandt mit dem Data-Mining ist.[107] Das wird auf verschiedene Gründe zurückgeführt. Teils wird dahingehend abgegrenzt, dass das Data-Mining der Datenbeschaffung und -aufbereitung für die weitere Analyse durch KI-Algorithmen, z. B. im Bereich des maschinellen Lernens, diene, d. h. primär keine eigenen Erkenntniszwecke verfolgt werden.[108] Andere sehen Überschneidungen jedenfalls darin, dass die TDM-Daten zugleich für das maschinelle Lernen genutzt werden können.[109] Oft wird das Training von KI-Algorithmen zu der Technologie gezählt, die in dieser Arbeit als die TDM-Analyse im engeren Sinne bezeichnet wird.[110] Zudem kommt gerade beim Data-Mining häufig unüberwachtes Lernen zum Einsatz oder es wird ein Modell mittels maschinellem

105 *Ertel*, Grundkurs Künstliche Intelligenz, S. 164 ff.

106 *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 9 ff.; *Ertel*, Grundkurs Künstliche Intelligenz, S. 164 ff. erläutert diesbezüglich „*LEXMED*", einem System z. B. zur Ermittlung der Wahrscheinlichkeit einer Blinddarmentzündung beim Vorliegen bestimmter Blutwerte oder Symptomatiken zur Unterstützung der medizinischen Diagnostik herangezogen werden kann.

107 *Lenzen*, Künstliche Intelligenz, S. 50 f.; *Sag*, CSUSA 2019, S. 291, 299 f.

108 *Leimeister*, Einführung in die Wirtschaftsinformatik, S. 177 f., 179 f. Nicht ganz stringent ist seine Argumentation aber deswegen, weil er die benannten Analysezwecke im weitesten Sinne dennoch unter das Data-Mining fasst.

109 *Steinrötter/Schauer* in Barudi, UrhG, § 4 Rn. 5 mit Verweis auf BT-Drs. 19/27426, S. 60, die eine Bedeutung des TDM für die KI-Entwicklung betont; ähnlich auch *Spindler*, CR 2019, S. 277, 278 Rn. 2; *B. Raue*, ZUM 2019, S. 684, 684 f.

110 *Sag*, CSUSA 2019, S. 291, 291 f.; *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 2 f.; *Ertel*, Grundkurs Künstliche Intelligenz, S. 206 f.; *Grätz*, Künstliche Intelligenz im Urheberrecht, S. 27 f.; *Käde*, Kreative Maschinen und Urheberrecht, S. 58 f.; *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 3.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Lernen an einem kategorisierten Datenset trainiert, um am restlichen Datensatz unüberwacht zum Einsatz zu kommen.[111] Insofern basiert das Data-Mining oftmals auf maschinellem Lernen, das betonen z. B. *Görz, Schmid* und *Braun*: „*Verfahren des Maschinellen Lernens sind die Grundlage von Programmsystemen, die aus »Erfahrung« lernen, also neues Tatsachen- und Regelwissen gewinnen oder Priorisierungen adaptieren können. Sie sind u. a. auch für die Entdeckung zweckbestimmt relevanter Beziehungen in großen Datenmengen (Data-Mining) von großer Bedeutung*".[112] In die gleiche Richtung argumentiert *Ertel*: „*Der Prozess des Gewinnens von Wissen aus Daten sowie dessen Darstellung und Anwendung wird als Data Mining bezeichnet. Die verwendeten Methoden kommen meist aus der Statistik oder dem maschinellen Lernen und sollten auch auf sehr große Datenmengen mit vertretbarem Aufwand anwendbar sein*".[113] Weiter bezeichnet er das Data-Mining aber auch als ein Teilgebiet der KI, das sich der Anforderung widmet, „*aus großen Datenbanken explizites Wissen zu gewinnen*".[114] Eine weitere Parallele bildet, dass auch für den KI-Trainingsprozess eine Aufbereitung der Trainingsdaten in unterschiedlichem Umfang erforderlich ist.[115]

Der Vergleich des TDMs mit der Wissensentdeckung in Datenbanken (KDD)[116] ist auf die besondere Fähigkeit von TDM-Verfahren zurückzuführen, Wissen aus großen Datensätzen zu extrahieren. Dieser Fähigkeit bedienen sich auch wissensbasierte Systeme und ihre spezielle Ausprägung der Expertensysteme. Der Zusammenhang von TDM und wissensbasierten Systemen bzw. Expertensystemen liegt also darin, dass die notwendige Wissensbasis mit TDM schneller und besser aufgebaut werden kann und aus der Datenbasis zuverlässiger, sorgfältiger und effizienter weitere Erkenntnisse extrahiert werden können, die Grundlage für die Wissensverarbeitung und die späteren Handlungsempfehlungen der Systeme sind.[117]

Aus den verschiedenen Differenzierungen sind die folgenden Erkenntnisse zu gewinnen: Technologien, die zur KI zählen, wie das maschinelle Lernen, Expertensysteme und das TDM, können aufeinander basieren. Das Auffinden von Strukturen und Zusammenhängen kann also das Ziel verfolgen, Algorithmen mit maschinellem Lernen zu trainieren oder eine

111 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 290 f.
112 *Görz/Braun/Schmid* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 1, 20.
113 *Ertel*, Grundkurs Künstliche Intelligenz, S. 206.
114 *Ertel*, Grundkurs Künstliche Intelligenz, S. 11.
115 Vgl. dazu *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 290 f.
116 S. dazu S. 30 ff.
117 *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 145.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Wissensbasis für Expertensysteme aufzubauen. Das Data-Mining dient daneben als eigenes Forschungswerkzeug, wie z. B. in den digitalen Geisteswissenschaften. Jedenfalls der Trainingsprozess im maschinellen Lernen ist als Data-Mining einzuordnen, denn dort ist das primäre Ziel, Muster und Strukturen aufzudecken. Diese haben lediglich nicht zwingend das Ziel, unmittelbare Erkenntnisse zu gewinnen. Zielführend erscheint es jedenfalls für diese Arbeit, unter der Anerkennung einer begrifflichen Unschärfe eine zweckorientierte Abgrenzung vorzunehmen, bei der das TDM als Oberbegriff fungiert und im Einzelfall weiter differenziert wird.

D. Ein Ablauf des Text- und Data-Minings

Im Beispiel von *Roberto Busa*[118] machten die Vorverarbeitungsschritte einen erheblichen Teil des Forschungsprozesses aus. Die Transferierung von Ursprungsmaterialien zu einem Forschungskorpus bildet auch heute den Abschnitt des TDM-Prozesses, der sich am kosten- und zeitaufwendigsten gestaltet und auf dem gleichzeitig die gesamte Forschungsarbeit im TDM aufbaut, denn eine gute Strukturierung und Aufbereitung der Forschungsdaten ist unerlässlich, um (belastbare) Analyseergebnisse zu erhalten. Die Vorverarbeitungsschritte betreffen ca. 70 bis 80 Prozent des Gesamtprozesses.[119] Gerade bei großen Datensätzen kommen zumeist schon zu diesem Zeitpunkt automatisierte Verfahren zum Einsatz, die die Vorverarbeitung vornehmen und die ihrerseits auf maschinellem Lernen basieren können.[120] Zu den relevanten Schritten zählen die Datenauswahl sowie die Vorverarbeitung der analysierten Texte und Daten wie die Digitalisierung, die Extrahierung des Reintextes aus Digitalisaten, die Segmentierung, die Bereinigung und die Annotation. Die vorverarbeiteten Daten werden sodann in das Korpus zusammengefügt, das anschließend Gegenstand der TDM-Analyse im engeren Sinne ist.

Im Folgenden wird ein Ablauf des Gesamtprozesses des TDMs dargestellt, der zwar keine Allgemeingültigkeit für sich beanspruchen kann, aber in unterschiedlichen Ausprägungen vorstellbar ist und der Untersuchung zugrunde gelegt wird. Das technische Verständnis bildet die Grundlage zur Beantwortung der Frage, was TDM-Methoden leisten können,

118 Vgl. dazu S. 36 f.
119 *Schaffranietz/F. Neumann* in: Keuper/Neumann (Hrsg.), Wissens- und Informationsmanagement, S. 149, 160; *Petersohn,* Data Mining, S. 40.
120 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 180 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

welches Potenzial sie haben und wo ihre Grenzen liegen. Weiter ist der technische Ablauf Maßstab der später folgenden urheberrechtlichen Betrachtung.[121]

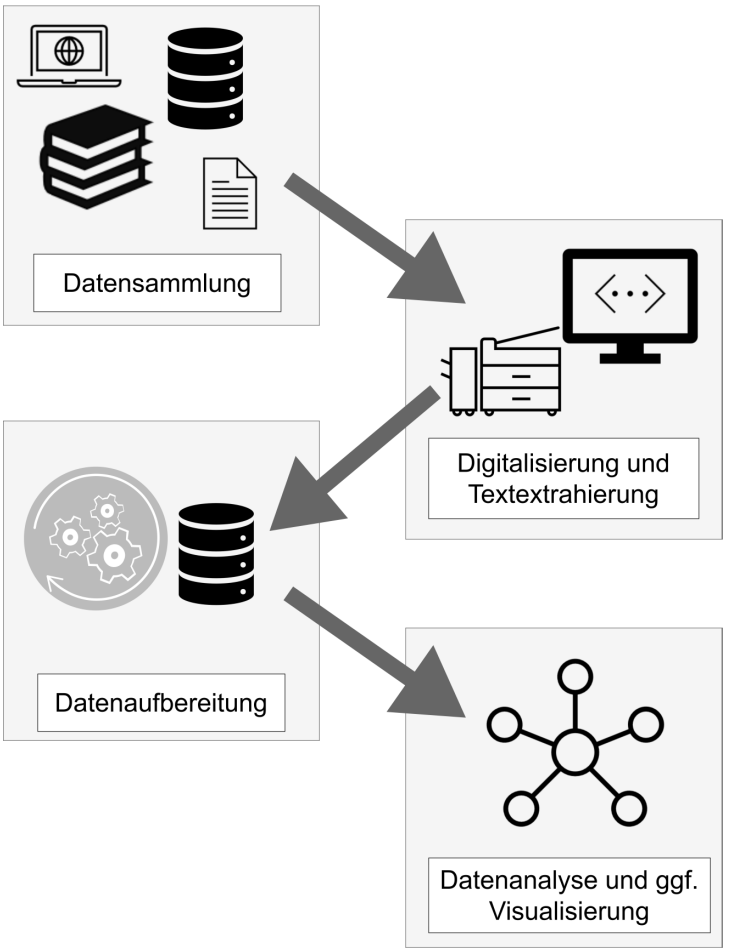

Abbildung 1: Gesamtprozess des TDMs (vereinfacht). Quelle: Eigene Darstellung.

121 Diese erfolgt im zweiten Teil der Arbeit.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Die in der Abbildung 1 zusammengefasst dargestellten und im folgenden Abschnitt erläuterten Schritte handeln sich dennoch nur um eine gewisse Bandbreite möglicher Forschungsabschnitte, die die Komplexität und den Umfang der Forschung veranschaulichen sollen.

Es kann jedoch nur ein exemplarischer Ablauf eines vorstellbaren Forschungsprozesses dargestellt werden: Die automatisierte digitale Daten- und Textanalyse ist noch nicht so etabliert, als dass sich ein Verfahren als das übliche bezeichnen ließe.[122] Dennoch zeichnen sich gerade in den digitalen Geisteswissenschaften nicht zuletzt aufgrund der jüngst starken Förderung derartiger Forschungsprojekte entsprechende Tendenzen ab.[123] Die Arbeitsschritte sind aufgrund der Abstraktion der folgenden Darstellung auch auf andere Anwendungen – auch auf das KI-Training – übertragbar, denn auch das maschinelle Lernen erfordert umfangreiche Vorverarbeitungsschritte,[124] die in die Datengewinnung, die Vorverarbeitung, die Datenspeicherung und die Datenanalyse aufteilbar sind.[125] Aus diesem Grund werden die Arbeitsschritte weitestgehend anwendungsneutral analysiert.

I. Datenauswahl und Selektierung

Bei den Ressourcen, die im TDM analysiert werden, kann es sich um Daten, Bilder, Audio-Dateien, Werte, Videos, etwa aus den sozialen Medien, von Unternehmen oder Daten der öffentlichen Hand handeln, die teils bereits angeordnet in Form von Datenbanken, teils auch unstrukturiert

122 Zu nennen ist allerdings CRISP-DM (Cross-Industrie-Standard-Process for Data Mining), ein standardisiertes Verfahren zum Data Mining, das im Jahr 1996 im Rahmen eines EU-Projekts entwickelt wurde. Es richtet sich aber vorwiegend an Wirtschaftsunternehmen und soll ihnen einen branchenoffenen Leitfaden bieten, *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 199 f.

123 Zu der Vorgehensweise in den Digital Humanities detailliert *Reiter/Pichler/Kuhn,* Reflektierte algorithmische Textanalyse (2020); *Jannidis/Kohle/Rehbein,* Digital Humanities; dagegen zur empirischen Sozialforschung detailliert *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103 ff.

124 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 179 f. bezeichnet diese Schritte als „*Pipeline*".

125 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 179 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

vorliegen.[126] Im Bereich des Text-Minings können sämtliche Texte Forschungsgegenstand sein, twa fiktionale Literatur, Sach- oder Fachliteratur oder Zeitungen. Für die Gewinnung von Daten aus dem Web können u. a. Web-Crawler genutzt werden.[127]

Um zu verhindern, dass Werke oder Schutzgegenstände für digitale Analysen durch automatisierte Verfahren gewonnen werden, werden häufig technische Barrieren eingesetzt. Dazu zählen sog. robot.txt-Dateien, das sind Textdateien, die von Web-Crawlern oder -Scrapern primär angefragt werden und die Angaben darüber enthalten, welche Inhalte nach Auffassung der Webseitenbetreiber abgegriffen werden dürfen und welche nicht,[128] sowie sog. CAPTCHA-Programme,[129] bei denen der Zugriff auf dynamische Webseiteninhalte vom Bestehen eines Tests abhängig gemacht wird, der etwa das Erkennen von Symbolen auf einem Bild verlangt.[130]

II. Die Vorverarbeitungsschritte

Die Vorverarbeitung der Forschungsdaten ist Grundlage der darauffolgenden TDM-Analyse im engeren Sinne sowie Voraussetzung der Funktionsfähigkeit und Leistungsfähigkeit von Technologien, die zur KI zählen. Das wird gerade dadurch deutlich, dass es aus technischer Sicht im Bereich des maschinellen Lernens sehr schwierig ist, maschinell getroffene Entscheidungen nachzuvollziehen[131] sowie beinahe unmöglich ist, sog. „*Machine Unlearning*“ zu betreiben, d. h. trainingsbedingte Fehler nachträglich zu

126 Ein Beispiel für solche strukturierten Daten sind auch sog. Data Warehouses, die vorwiegend auf unternehmerischer Ebene genutzt werden, *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 286.

127 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 180.

128 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 27; *v. Schönfeld,* Screen Scraping und Informationsfreiheit, S. 60; *Schrenk,* Webbots, spiders, and screen scrapers, S. 311.

129 CAPTCHA ist eine Abkürzung für „*Completely Automated Public Turing test to tell Computers and Humans Apart*“, *v. Schönfeld,* Screen Scraping und Informationsfreiheit, S. 64; *Schrenk,* Webbots, spiders, and screen scrapers, S. 314.

130 *A. Schmidt,* Virtuelles Hausrecht und Webrobots, S. 33.

131 Die besondere Schwierigkeit liegt darin, dass nicht ganz nachvollzogen werden kann, welcher Datenpunkt welchen Einfluss auf den trainierten Algorithmus hat, *Bourtoule/Chandrasekaran et al.*, Proceedings of the 42nd IEEE Symposium on Security and Privacy. Diese Black-Box wird auf S. 56 ff. näher besprochen.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

beheben.[132] Kleine Abweichungen oder Fehler können infolgedessen zu einer erheblichen Verzerrung oder Verfälschung der Analyseergebnisse führen.[133] Das kann sich ebenso negativ auf die darauffolgenden Forschungen bzw. auf die Nutzung trainierter Systeme auswirken.

Im Text-Mining und im Data-Mining stellen sich die Schwerpunkte der Vorverarbeitungsschritte häufig unterschiedlich dar. Die Vorverarbeitungsschritte im Text-Mining dienen vorwiegend dazu, die Datensätze maschinenlesbar und durchsuchbar zu machen sowie sie für die spätere Analyse vorzubereiten, sodass Zusammenhänge oder andere Erkenntnisse herausgefiltert werden können. Die Vorverarbeitung im Data-Mining hat den wesentlichen Unterschied, dass die Daten bereits digital und maschinenlesbar vorhanden sind. Die Vorverarbeitung im Data-Mining dient insofern vorwiegend dazu, die Daten zu harmonisieren, Fehler in den Datensätzen zu korrigieren bzw. die Daten zu bereinigen, um falsche Ergebnisse zu vermeiden und die Daten auf diejenigen einzugrenzen, die für die Forschungsfrage relevant sind.[134]

1. Digitalisierung und Textextrahierung

Je nach Datenauswahl und Analyseziel kann es notwendig sein, im ersten Schritt analoge Quellen zu digitalisieren. Das Ergebnis einer einfachen Digitalisierung, also eines Scans, gibt den Text in Form eines Bildes wieder, z. B. im PDF- oder im JPEG-Format. Um diese Darstellungsformen maschinenlesbar zu machen, muss im nächsten Schritt der Reintext extrahiert

---

132 *Bourtoule/Chandrasekaran et al.*, Proceedings of the 42nd IEEE Symposium on Security and Privacy bezogen sich dabei auf den Datenschutz, der die nachträgliche Löschung erforderlich machen kann. Sie schlugen diesbezüglich vor, das spätere Machine Unlearning bereits von Anfang an zu berücksichtigen, indem Trainingssets zu Beginn aufgeteilt werden und die Algorithmen isoliert an diesen Teilsets trainiert werden, sodass kleinteilig nachvollzogen werden kann, ab welchem Punkt ein Datensatz Einfluss auf das Trainingsergebnis hatte.

133 *J. Kuhn* in: Reiter/Pichler/Kuhn (Hrsg.), Reflektierte algorithmische Textanalyse, S. 9, 18; *Eckart/Gärtner/J. Kuhn/K. Schweitzer* in: Witt/Lobin/Schneider (Hrsg.), Digitale Infrastrukturen für die germanistische Forschung, S. 115, 117; *Fayyad/Piatetsky-Shapiro/Smyth* in: Fayyad (Hrsg.), Advances in knowledge discovery and data mining, S. 1, 4; *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 39.

134 *Bramer*, Principles of Data Mining, S. 13 f.; in Bezug auf die Vorverarbeitung von KI-Trainingsdaten *Niederée/Nejdl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 104 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

werden.[135] Den regulären Digitalisierungen, d. h. den einfachen Scans, die nur fotografische Abbilder einer Seite darstellen, werden dazu maschinenlesbare Zeichen extrahiert. Das wird als Optische Zeichenerkennung (englisch: *Optical Character Recognition* (OCR)) bezeichnet. Dabei wird die Sprache in Pixelanhäufungen erkannt und aus ihr extrahiert.[136] In einem weiteren Schritt wird der Text in der Regel in einen Reintext (z. B. in ein einfaches Textdokument im txt-Format) überführt.[137]

2. Segmentierung

Im Anschluss werden die Texte maschinell in kleine Segmente, sog. Tokens, umgewandelt.[138] Diese Tokens sind dabei diejenigen Elemente eines Texts, die als kleinste Ebene der Annotation und Kategorisierung festgelegt werden, beispielsweise Worte, Haupt- oder Nebensätze oder auch Zeichen, Silben oder Laute.[139] Zumeist werden die Tokens in Tabellenstruktur untereinander angeordnet, z. B. im XML-Format.[140]

135 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 193.

136 Dies basiert heute auch auf Methoden der KI, die im Bereich der Spracherkennung leistungsfähiger ist, als herkömmliche Software, S. dazu bereits S. 22 ff.; *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 193.

137 *Jo*, Text Mining, S. 6.

138 *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1106 f. Der Begriff „*Token*" stammt aus der Computerlinguistik und soll Missverständnisse oder Streitigkeiten dazu vermeiden, welche Satzelemente als Wörter bezeichnet werden und welche nicht, *Carstensen/Ebert et al.*, Computerlinguistik und Sprachtechnologie, S. 264 f.; *Petz*, Opinion Mining im Web 2.0, S. 169; *Jo*, Text Mining, S. 5, 19, 21; *Zong/Xia/Zhang*, Text Data Mining, S. 20.

139 *Hirschmann*, Korpuslinguistik, S. 31 f.; *Jo*, Text Mining, S. 5, 19, 21 dagegen versteht unter Tokens die Elemente, die durch ein Leer- oder Satzeichen voneinander getrennt sind.

140 XML ist die Abkürzung für die Extensible Markup Language, die ist ein Dateiformat ist, das derzeit als Standardformat zur Beschreibung textueller Datenelemente gilt, *Jo*, Text Mining, S. 6.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

3. Bereinigung

Insbesondere Texte werden gelegentlich um sog. Stopp-Wörter bereinigt.[141] Das kann etwa in Betracht kommen, wenn bestimmten Wörtern oder Segmenten nur eine geringe oder keine semantische Bedeutung beigemessen wird und sie deswegen nicht für das Analyseergebnis relevant sind[142] oder darauf sogar negative Einflüsse haben, indem sie es verzerren oder verfälschen.[143] Bei Daten, die mit Web-Crawlern gewonnen wurden, betrifft das z. B. auch Elemente, die nicht den Inhalt, sondern nur die Darstellung betreffen.[144] Die möglichen Stoppwörter variieren je nach Forschungsmaterial und -frage, es kann sich etwa um direkte Zitate,[145] um Artikel, Pronomen oder Präpositionen handeln.[146] Gerade in den digitalen Literaturwissenschaften als Teilgebiet der digitalen Geisteswissenschaften sind es typischerweise Ortsangaben, Namen oder Eigennamen.[147]

4. Metadaten und Annotationen

Ein für die Beforschung einzelner Forschungsfragen und für das Training von KI-Algorithmen zentraler Schritt ist die manuelle oder automatisierte Annotation, bei der weitere Informationen an die segmentierten Datensammlungen annotiert, d. h. angefügt, werden. Oftmals ermöglichen erst diese weiteren Informationen, dass entweder KI-Systeme lernen können oder dass algorithmische automatisierte Verfahren Aussagen über die Datenbestände treffen bzw. Erkenntnisse über die Datensammlungen offen-

---

141 Zur Identifizierung von derartigen Wörtern gibt es Datenbanken, je nach Forschungsfrage müssen die Stopp-Wörter aber dennoch selbst definiert werden, vgl. *Bramer,* Principles of Data Mining, S. 332; *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1107; *Zong/Xia/Zhang,* Text Data Mining, S. 20.

142 *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1107.

143 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 181.

144 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 180 f.

145 *Kleinkopf,* Text- und Data-Mining in den digitalen Geisteswissenschaften.

146 *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1107.

147 *Horstmann,* forTEXT. Literatur digital erforschen Rn. 15; *Schuhmacher,* forTEXT. Literatur digital erforschen Rn. 4 ff. Ein Verfahren, mit dem diese Stoppwörter individuell ausgewählt und automatisiert herausgelöscht werden können, ist die sog. *Named Entity Recognition.*

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

baren können. Die Annotationen können an einzelne Tokens angefügt werden oder sich auf größere Segmente des Korpus beziehen, z. B. auf diejenigen Abschnitte, die aus derselben Quelle stammen.

Die Abgrenzung von Annotationen und Metadaten ist in den Fachdisziplinen, die Korpora beforschen, hoch umstritten.[148] Unter Metadaten werden in dieser Arbeit diejenigen Informationen verstanden, die sich auf größere Korpuspassagen beziehen und als Annotationen diejenigen Informationen, die an einzelne Tokens annotiert werden. Vorstellbare automatisierte Annotationen sind u. a. Wortarten, Angaben zum Satzbau oder zur Formatierung der Texte[149] sowie die Lemmata, d. h. die Wortstämme (sog. Lemmatisierung, englisch: Stemming).[150] Automatisierte Verfahren zur Annotation und zur Anfügung von Metadaten können dabei freilich Zeit sparen. Insbesondere in den digitalen Geisteswissenschaften werden Annotationen jedoch jedenfalls teilweise manuell angefügt, z. B. zur Entwicklung eines sog. Goldstandards zum Training neuronaler Netze, um weitere Annotationen anhand dieses Goldstandards zu automatisieren.[151] Die Metadaten und Annotationen bilden eine Kategorisierung, d. h. eine

148 Z. B. in der digitalen Literaturwissenschaften wird derart differenziert, dass die Annotationen die einzelnen Segmente (Tokens) betreffen, die Metadaten größere Zusammenhänge, *Bläß,* forTEXT. Literatur digital erforschen Rn. 1. In der Korpuslinguistik wird wiederum nur als Annotation eingeordnet, was linguistisch kategorisiert, während unter Metadaten ebenso die sekundären Informationen über die Korpora oder Teile der Korpora verstanden werden, *Hirschmann,* Korpuslinguistik, S. 2. Letztere Ansicht lässt sich nur schwer auf andere Forschungsbereiche übertragen, deswegen erscheint eine Abgrenzung dahingehend, dass sich Metadaten auf das Gesamtkorpus oder Korpusteile beziehen, während Annotationen sich auf kleinere Textsegmente, z. B. Tokens, beziehen, zielführender.

149 *Eckart/Gärtner/J. Kuhn/K. Schweitzer* in: Witt/Lobin/Schneider (Hrsg.), Digitale Infrastrukturen für die germanistische Forschung, S. 115, 120. Für automatisierte Annotationen greifen Annotationswerkzeuge häufig auf digitale Lexika zurück, vgl. *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1107.

150 *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1107; *Hirschmann,* Korpuslinguistik, S. 37; im Englischen wird dies als Stemming bezeichnet, vgl. *Bramer,* Principles of Data Mining, S. 332; *Jo,* Text Mining, S. 19 f.

151 In diesem Zusammenhang kommt auch gelegentlich die Methode zum Einsatz, in Web-Anwendungen auf Basis eines Spiels möglichst viele Nutzer zu erreichen, die für die Vornahme einzelner Annotationen kleine Belohnungen erhalten (sog. Gamification), *Jannidis/Kohle/Rehbein,* Digital Humanities, S. 260, 290 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Einordnung der Tokens und helfen der Software dabei, zu lernen.[152] Insofern bilden die Annotationen einen elementaren Pfeiler des überwachten Lernens.[153] Die folgende Abbildung zeigt einen annotierten Korpusauszug exemplarisch anhand des zweiten Satz des achten Erwägungsgrundes der DSM-RL, um die erläuterten Annotationen am Beispiel linguistischer Annotationen zu veranschaulichen. Vorstellbar wäre in diesem Fall etwa, diese Daten als Teil eines größeren Datensets, das ebenso aus der DSM-RL oder aus anderen europäischen Richtlinien stammt, in Bezug auf die Eigenheiten der Sprache europäischer Richtlinientexte zu analysieren oder ein Modell mittels maschinellen Lernens auf diese Sprache zu trainieren. Mittels darüber hinausgehenden inhaltlichen Annotationen könnten auch semantische Zusammenhänge extrahiert werden.

```
1       Mittels _  Präposition _ Objekt | Kasus=Genitiv
2       Text- und Data-Mining _ Substantiv _  Objekt | Kasus=Genitiv | Numerus=3PSingular | Genus=neutr
3       lassen _  Verb _ Prädikat | Tempus=Präsens | Form=Inf Numerus=3PPlural | Genus verbi=passiv
4       sich    _ Pronomen _ Prädikat | Tempus=Präsens | Form=Inf Numerus=3PPlural | Genus verbi=passiv
5       große    _ Adjektiv _ Subjekt | Numerus=3PPlural
6       Informationsmengen _ Substantiv _ Subjekt | Numerus=3PPlural
7       verarbeiten _ Verb _ Prädikat | Tempus=Präsens | Form=Inf | Numerus=3PPlural | Genus verbi=passiv
8       ,
9       um _ Konjunktion _ Deutsch
10      neue _ Adjektiv _ Objekt | Kasus=Akkusativ | Numerus=3PPlural | Genus=fem
11      Erkenntnisse _ Substantiv _ Objekt | Fall=Akkusativ | Numerus=3PPlural | Genus=fem
12      zu gewinnen _ Verb _ Prädikat | Tempus=Präsens | Form=Inf
13      und _ Konjunktion _ Deutsch
14      neue _ Adjektiv _ Objekt | Kasus=Akkusativ | Numerus=3PPlural | Genus=mask
15      Trends _ Substantiv _ Objekt | Kasus=Akkusativ | Numerus=3PPlural | Genus=mask
16      zu erkennen _ Verb _ Prädikat | Tempus=Präsens | Form=Inf
17      .
```

Abbildung 2: Linguistische Annotation eines Beispielssatzes aus der DSM-RL. Quelle: Eigene Darstellung.

III. Das Korpus

Die Zusammenstellung der aufbereiteten Daten bzw. Textdaten, z. B. in XML-Tabellenform, wird als das Korpus bezeichnet. Der Begriff des Korpus wird zwar vorwiegend in den digitalen Geisteswissenschaften genutzt, ist aber genauso für KI-Trainingsdaten anwendbar, denn er steht schlicht für die Sammlung der aufbereiteten Daten, die im sich daran anschließenden Schritt algorithmusbasiert ausgewertet wird.

Aus Perspektive der digitalen Geisteswissenschaften werden Korpora als *„digitale Sammlungen von Text oder gesprochener Sprache“* bezeichnet, *„die*

152 *Truyens/van Eecke*, CLSR 2014, S. 153, 154; *Jannidis/Kohle/Rehbein,* Digital Humanities, S. 260.

153 *Truyens/van Eecke*, CLSR 2014, S. 153, 154.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

linguistisch annotiert wurden sowie (in vielen Fällen) für eine bestimmte Domäne repräsentativ sind und somit als Referenz fungieren können".[154] Klarstellend sollen die Ursprungsdaten selbst nicht erfasst sein, sondern nur ihre digitalen und ggf. auch sonst verarbeiteten Repräsentationen.[155]

Die Korpuslinguistik, die sich im Kernbereich mit der Untersuchung linguistischer Korpora befasst,[156] die mit digitalen Mitteln beforscht werden, ist sich uneinig darüber, was Korpora eigentlich sind.[157] Nach einem engen Verständnis ist jedenfalls ein linguistisches Korpus eine digitale Textsammlung, die von Sprachwissenschaftlern für die Zwecke sprachwissenschaftlicher Analysen erstellt wurden.[158]

Aus literaturwissenschaftlicher Perspektive wird das Korpus als „*maschinenlesbare Sammlung von Texten, die den Gegenstand*" der „*Untersuchungen im Feld digitaler Literaturwissenschaft bildet*" betrachtet.[159] Anknüpfungspunkt ist also, dass das Korpus der Analysegegenstand und als solcher individuell konzipiert ist sowie, dass er durch eine Vorverarbeitung entstand.[160]

Begrifflich ist das Korpus dem Juristen aus dem rechtsgeschichtlichen Kontext bekannt, dem Corpus Iuris Civilis, einer Gesetzessammlung aus dem Römischen Reich. Insofern findet sich eine Parallele in Bezug auf den Charakter der Zusammenstellung bzw. Sammlung verschiedener Inhalte.

Nach dem Duden ist ein Korpus eine „*Belegsammlung [antiker oder mittelalterlicher] von Texten oder Schriften*" oder eine „*als Datenbank angelegte Sammlung einer begrenzten Anzahl an Texten, Äußerungen o. Ä. als Grundlage für sprachwissenschaftliche Untersuchungen*".[161]

Davon abgesehen, dass die dargestellten Definitionen den Anwendungsbereich von Textsammlungen stark eingrenzen, indem sie sich nur auf

---

154 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 223; *McEnery/Wilson*, Corpus linguistics, S. 29 ff.

155 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 223.

156 Die Korpuslinguistik ist eine empirische Methode, sprachwissenschaftliche Fragestellungen zu untersuchen, *Hirschmann*, Korpuslinguistik, S. 1.

157 *Mair* in: Kupietz/Schmidt (Hrsg.), Korpuslinguistik, S. 5, 23: „*In gewisser Weise ist der Korpuslinguistik sogar ein zentrales Identifikationsmerkmal abhandengekommen, nämlich der Konsens darüber, was ein Korpus ist.*"; *Gatto*, ICAME Journal 2011, S. 35, 53.

158 *Mair* in: Kupietz/Schmidt (Hrsg.), Korpuslinguistik, S. 5, 23; *Hirschmann*, Korpuslinguistik, S. 2.

159 *Bläß*, forTEXT. Literatur digital erforschen Rn. 1.

160 *Bläß*, forTEXT. Literatur digital erforschen Rn. 1, 3.

161 So der Duden, Definition des Korpus (Sammlung), abrufbar unter: https://www.duden.de/rechtschreibung/Korpus_Sammlung.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

sprachwissenschaftliche Untersuchungen beziehen, bringen sie gut zum Ausdruck, dass das Korpus ein Datensatz ist, der Repräsentationen von Text oder Textextrakten, Zahlen und sonstiger Daten ist, die einem gewissen Thema angehören oder in einem sonstigen Zusammenhang stehen und zu einem bestimmten Zweck – z. B. der Datenanalyse – erstellt wurden.

IV. Die Text- und Data-Mining-Analyse im engeren Sinne

Die TDM-Analyse im engeren Sinne ist zwar nicht der zeitintensivste Abschnitt des Gesamtprozesses, aber doch das angestrebte Ziel. Die Korpora werden mittlerweile bevorzugt mittels Deep Learning und anderen Methoden aus dem Bereich des maschinellen Lernens analysiert, es werden aber auch statistische Verfahren eingesetzt,[162] wobei die Möglichkeiten bzw. das Potenzial solcher statistischer Verfahren in Relation zu Verfahren, die zum maschinellen Lernen zählen, stark begrenzt sind.[163]

Die folgende Darstellung der Analysemethoden erhebt keinen Anspruch auf Vollständigkeit, sie soll vielmehr veranschaulichen, auf welcher Grundlage die teils eigens programmierten Analyse-Tools lernen und entscheiden. Dabei unterscheiden sich die Anwendungsbereiche der verschiedenen Methoden selbstverständlich in Bezug auf die Fachdisziplinen, die beforschten Ressourcen sowie die individuellen Forschungsfragen oder Zielsetzungen, das gilt z. B. für die Frage, ob Methoden des überwachten oder unüberwachten Lernens eingesetzt werden. Das hängt davon ab, ob bereits definierte und z. B. über Annotationen und Metadaten festgelegte Kategorisierungen untersucht werden sollen (überwachtes Lernen), oder ob gerade interessant ist, welche Verbindungen ein Computer zwischen den jeweiligen Elementen herstellt, an die ein Mensch möglicherweise nicht gedacht hätte (unüberwachtes Lernen). Die dargestellten Methoden können genauso zum Ziel des Trainings von KI-Algorithmen, also im maschinellen Lernen zum Einsatz kommen.[164]

Eine verbreitete Methode der TDM-Analyse ist die *Klassifikation*, die zum überwachten Lernen gezählt wird.[165] Anhand bereits vorgenomme-

162 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 289.
163 *Jannidis/Kohle/Rehbein*, Digital Humanities, S. 289.
164 *Zong/Xia/Zhang*, Text Data Mining, S. 102, 111 ff.
165 *Bramer*, Principles of Data Mining, S. 6; *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 287; *Jo*, Text Mining, S. 8; *Leimeister*, Einführung in die

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

ner Kategorisierungen, d. h. den Annotationen und Metadaten, erlernt der Algorithmus dabei, auf welchen Gemeinsamkeiten die gleiche oder auf welchen Unterschieden die unterschiedliche Kategorisierung fußt, um im weiteren Schritt unbekannte Fälle einzuordnen.[166] Das Ergebnis der Klassifikation ist eine Wahrscheinlichkeit der Kategorisierung.[167] In den digitalen Geisteswissenschaften wird die Klassifikation genutzt, um bestimmte Inhalte bzw. inhaltliche Übereinstimmungen in größeren Textbeständen aufzufinden.[168] Dabei kann z. B. die Autorenschaft anhand der Stilistik festgestellt werden.[169] Bei der *Sentimentanalyse* werden mittels der Klassifikation Stimmungen analysiert.[170] Ziel ist, Informationen in größeren Textzusammenhängen zu finden und zu extrahieren, die menschliche Wahrnehmungen oder Meinungen betreffen.[171] Zugrunde liegen dabei Wortlisten, die eine Bewertung einzelner Worte vornehmen, z. B. als „*positiv*" oder „*negativ*".[172] Ein weiteres Anwendungsbeispiel sind Spam-Filter.[173]

Mit der *Regression*, die zum unüberwachten Lernen zählt, werden statt Kategorien numerische Werte prognostiziert, die Assoziationen abbilden.[174] Ein prominentes Beispiel hierfür sind Warenkorbanalysen, bei denen anhand des bisherigen Kaufverhaltens Wahrscheinlichkeiten für den Kauf anderer Produkte errechnet und somit auf das weitere Kaufverhalten geschlossen wird.[175]

Wirtschaftsinformatik, S. 187; *Zong/Xia/Zhang,* Text Data Mining, S. 3; *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 44. Genauer kann sie z. B auf dem „*Naive Bayes-*" oder dem sog. „*k-Nearest Neighbours*"-Verfahren basieren, *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 187; *Zong/Xia/Zhang,* Text Data Mining, S. 103.

166 *Bramer,* Principles of Data Mining, S. 6; *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 287 f.; *Jo,* Text Mining, S. 7; *Jannidis/Kohle/Rehbein,* Digital Humanities, S. 289; *Stiemerling*, CR 2015, S. 762, 763.

167 *Leimeister,* Einführung in die Wirtschaftsinformatik, S. 191.

168 *Schöch/Döhl et al*., ZfdG 2020 Rn. 28.

169 *Schöch/Döhl et al*., ZfdG 2020 Rn. 28.

170 *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1113; *Jo,* Text Mining, S. 96; *Schöch/Döhl et al*., ZfdG 2020 Rn. 35.

171 *Flüh,* forTEXT. Literatur digital erforschen Rn. 1 f.

172 *Flüh,* forTEXT. Literatur digital erforschen Rn. 26.

173 *Jo,* Text Mining, S. 94 f.

174 *Bramer,* Principles of Data Mining, S. 6; *Jo,* Text Mining, S. 9; *Fayyad/Piatetsky-Shapiro/Smyth*, AI Mag 1996, S. 37, 44.

175 *Bramer,* Principles of Data Mining, S. 8.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Eine weitere Methode, die (zumeist) im unüberwachten Lernen in TDM-Analysen zum Einsatz kommt, ist die *Cluster-Analyse* bzw. das *Clustering*,[176] mithilfe dessen Ähnlichkeiten extrahiert und daraus Gruppen, die Cluster, gebildet werden.[177] Infolgedessen kann auch das *Clustering* dazu verwendet werden, inhaltliche Gemeinsamkeiten in größeren Datensätzen zu orten.[178] Mittels des *Topic Modeling* werden des Weiteren unüberwacht statistische Häufungen mehrerer Wörter erfasst, um unterschwellige Zusammenhänge in größeren Text- bzw. Datensammlungen aufzufinden, in der digitalen Literaturwissenschaft betrifft das z. B. abstrakte Themen oder Motive.[179]

V. Das Analyseergebnis – Einordnung, Potenzial und Validität

Die Ausgabe der TDM-Analyse im engeren Sinne kann sehr unterschiedlich ausfallen. Wie jede Forschungsmethode bedarf auch das Ergebnis der TDM-Analyse einer wissenschaftlichen Überprüfung, häufig auch bezüglich ihrer wissenschaftlichen Relevanz.[180] Die tatsächliche Einordnung der Ergebnisse fällt etwa im Fall der digitalen Geisteswissenschaften gerade deswegen schwer, weil der Forschungsbereich noch sehr jung ist und die Prozesse in weiten Teilen stark experimentell sind.[181]

Nach dem gängigen Verständnis in der informationstechnischen Literatur bringt das TDM *„neues Wissen"* zutage, indem es versteckte und dem Menschen nicht zugängliche Muster oder Strukturen offenbart.[182] Die ver-

176 *Leimeister*, Einführung in die Wirtschaftsinformatik, S. 184.; *Zong/Xia/Zhang*, Text Data Mining, S. 3.

177 *Leimeister*, Einführung in die Wirtschaftsinformatik, S. 184.; *Bramer*, Principles of Data Mining, S. 8; *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 287; *Jo*, Text Mining, S. 9; *Leimeister*, Einführung in die Wirtschaftsinformatik, S. 184.

178 *Schöch/Döhl et al*., ZfdG 2020 Rn. 28.

179 *Horstmann*, forTEXT. Literatur digital erforschen Rn. 1 f., 11 ff.; *Zong/Xia/Zhang*, Text Data Mining, S. 4; *Schöch/Döhl et al*., ZfdG 2020 Rn. 31.

180 *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 286.

181 *Döhl*, RuZ 2020, S. 195, 206: *„Der ganze Prozess ist immer noch vom Basteln und Ausprobieren gekennzeichnet, nicht zuletzt, weil gerade in den letzten ein bis zwei Jahren die digitalen Analysemethoden für Texte noch einmal erheblich verbessert wurden. Mit welchen digitalen Methoden welche interessanten Ergebnisse für die Geistes- und Kulturwissenschaften ermittelt werden können, ist eine weitgehend offene Frage"*, *Jannidis*, ZMK 2019, S. 63, 70.

182 *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 285 f.; *Beierle/Kern-Isberner*, Methoden wissensbasierter Systeme, S. 146 zum KDD, das dem hier vertretenen

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

schiedenen Gesetzes- und Richtlinienmaterialien gehen davon aus, dass die Analyseergebnisse „*Muster, Trends und Korrelationen*", Art. 2 Abs. 2 DSM-RL oder „*statistische Häufigkeiten oder Korrelationen*"[183] seien. Der rechtswissenschaftlichen Literatur liegt dasselbe Verständnis zugrunde, dort werden die Analyseergebnisse als „*Zusammenhänge*",[184] oder „*statistische Häufigkeiten oder Korrelationen*" eingeordnet.[185]

„*Muster*" können im wörtlichen Sinne grafische und sich wiederholende Strukturen meinen oder sich im übertragenen Sinne auf inhaltliche Zusammenhänge oder Referenzen in Textmaterialien beziehen, sie können für Vernetzung oder für Regelmäßigkeiten stehen und der Sprache, Bildern und der Gesellschaft[186] immanent sein. In der Forschung mit TDM wird in den analysierten Korpora nach inhaltlichen oder formalen Regelmäßigkeiten gesucht. Um Muster erkennen zu können, ist es erforderlich, verschiedene Merkmale zu extrahieren, die Untersuchungsobjekte aus verschiedensten Blickwinkeln zu betrachten und die Merkmale zueinander in Bezug zu setzen.[187] Der Begriff der Muster impliziert also sehr treffend, was das Ergebnis einer TDM-Analyse im engeren Sinne ist und wie unterschwellig und zugleich vielfältig die aufgefundenen Ergebnisse sein können.

Es ist nicht nur dem Wesen des Menschen,[188] sondern insbesondere der Wissenschaft immanent, Erklärungen, Ursachen und Antworten auf Fragen nach dem „*Warum*" zu suchen.[189] Ob dies auch im Rahmen von

---

Verständnis des TDM entspricht: „*Genauer versteht man unter Knowledge Discovery den Prozess, neues, nützliches und interessantes Wissen aus Daten herauszufiltern und in verständlicher Form zu präsentieren*"; *Ertel,* Grundkurs Künstliche Intelligenz, S. 11, 206: „*Der Prozess des Gewinnens von Wissen aus Daten sowie dessen Darstellung und Anwendung wird als Data Mining bezeichnet*"; kritisch bzgl. der Begriffsverwendung von Daten, Wissen und Information *W. Zimmer,* Ansturm der Algorithmen, S. 196.

183 BT-Drs. 18/12329, S. 40.

184 *de la Durantaye,* Allgemeine Bildungs- und Wissenschaftsschranke, S. 8; *B. Raue*, CR 2017, S. 656, 656.

185 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 44b Rn. 1; *Dreier* in Dreier/Schulze, UrhG, 6. Aufl. 2018, § 60d Rn. 1; *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60d Rn. 1; *Bullinger* in Wandtke/Bullinger, UrhG, 5. Aufl. 2019, § 60d Rn. 1.

186 Dazu stellte *Nassehi* etwa die Theorie auf, die Digitalität sei seit jeher ein dem System der Gesellschaft immanentes Muster, *Nassehi,* Muster.

187 *Nassehi,* Muster, S. 229 f.

188 *Pearl/Mackenzie,* The book of why, S. 1 ff.

189 Grundlegend zu wissenschaftlichen Erklärungen *Hempel/Oppenheim*, Philo. Sci. 1948, S. 135 ff.; *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

TDM-Forschungen technisch möglich ist, wird nachfolgend diskutiert. Anhand dieser Erwägungen wird eingeordnet, wie valide TDM-Analysen sind und was sie leisten können. Besondere Schwierigkeiten ergeben auch die bereits erwähnte schwierige Nachvollziehbarkeit der Prozesse des maschinellen Lernens.

1. Muster – Korrelation oder Kausalität?

Der Begriff der Kausalität stammt aus dem Lateinischen und bedeutet „*Ursächlichkeit*".[190] Die Bezeichnung der Korrelation geht hingegen auf das Wort „*relatio*" zurück, was in diesem Kontext mit einer „*wechselseitigen Beziehung*" sowie einem „*Wahrscheinlichkeitszusammenhang*" bzw. einem „*statistischen Zusammenhang*" gleichzustellen ist.[191] Eine Korrelation verlangt insofern einen statistischen Zusammenhang, eine Kausalität hingegen eine Ursache-Wirkungs-Beziehung.[192] Die tatsächliche wissenschaftliche Einordnung von einer Bedingung für die andere als kausal erfordert eine zeitliche Abfolge, eine statistische Korrelation sowie eine entsprechende Interpretation der Bedingungen als kausal.[193] Der statistische Zusammenhang, den eine Korrelation voraussetzt, kann hingegen in dritten Faktoren liegen, die ihrerseits auf die beiden untersuchten Phänomene wirken.[194] Dennoch erhöht sich die Wahrscheinlichkeit des Auftretens des einen Phänomens durch das Auftreten des anderen. Die Beziehung zwischen

---

und analytischen Philosophie, S. 110 ff.; *Kornmesser/Büttemeyer*, Wissenschaftstheorie, S. 135 ff.

190 So der Duden sowie das digitale Wörterbuch der deutschen Sprache zur Definition des Begriffs der Kausalität, abrufbar unter: https://www.duden.de/rechtschreibung/Kausalitaet; sowie unter: https://www.dwds.de/wb/Kausalit%C3%A4t.

191 So der Duden, Definition der Korrelation, abrufbar unter: https://www.duden.de/rechtschreibung/Korrelation; https://www.dwds.de/wb/Kausalit%C3%A4tM.

192 Nach mathematischem Verständnis wird unter einer Korrelation ein statistischer Zusammenhang zwischen verschiedenen Phänomenen verstanden; *Henze*, Stochastik für Einsteiger, S. 166 ff.; *W. Zimmer*, Ansturm der Algorithmen, S. 205. Den empirischen Wissenschaften kommt es bei einer Kausalität auf eine „*Beziehung zwischen Variablen*", auf eine Wirkung der einen auf die andere, an, *Kühnel/Dingelstedt* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1401, 1403.

193 *Weiber/Mühlhaus* in: Weiber/Mühlhaus (Hrsg.), Strukturgleichungsmodellierung, S. 9, 12.

194 Sog. Hintergrundvariablen, *Henze*, Stochastik für Einsteiger, S. 174.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

zwei korrelierenden Faktoren ermöglicht also die Abgrenzung von Korrelationen und Zufall.

Aus wissenschaftstheoretischer Perspektive ist das Ziel der wissenschaftlichen Forschung stets die Erklärung,[195] die zwingend die Antwort auf eine „*Warum*"-Frage enthält.[196] Zum Aufbau einer Erklärung sind wissenschaftstheoretisch zwei Modelle zu nennen, zum einen das deduktiv-nomologische, zum anderen das statistische Erklärungsmodell. Beide setzen voraus, dass eine Erklärung aus zwei Teilen besteht, zum einen aus dem Explanan, dem erklärenden Teil, und zum anderen aus dem Explanandum, dem zu erklärenden Teil,[197] wobei das Explanan naturwissenschaftliche Gesetze sowie sog. Antezedensbedingungen enthält,[198] wohingegen das Explanandum die Schlussfolgerung aus dem Explanan bildet.[199]

Im Rahmen des deduktiv-nomologischen Erklärungsmodells wird an die Verknüpfung des Explanans und des Explanandums das Erfordernis der deduktiven Gültigkeit[200] gestellt. Das bedeutet, dass es sich bei dem Explanandum um die logische Schlussfolgerung aus dem Explanan handeln muss[201] und beinhaltet, dass aus dem (Nicht-)Vorliegen des einen auch das (Nicht-)Vorliegen des anderen zu schlussfolgern ist.[202] Die Schlussfolgerung des Explanandums muss darüber hinaus kausal auf dem Explanan basieren.[203]

---

195 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 135.

196 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 135.

197 *Hempel/Oppenheim*, Philo. Sci. 1948, S. 135, 136 f.; *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 135; *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 123 f.

198 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 135 f.

199 In der juristischen Dogmatik könnte das Explanan etwa mit Definition und Subsumtion, das Explanandum mit der Schlussfolgerung oder dem Ergebnis gleichgesetzt werden.

200 Hempel und Oppenheim sahen natürlich auch, dass nicht jede Erklärung logisch schlussfolgernd erfolgen kann, *Wiltsche,* Einführung in die Wissenschaftstheorie, S. 186.

201 *Wiltsche,* Einführung in die Wissenschaftstheorie, S. 187.

202 Vorausgesetzt werden genauer die folgenden Punkte: Die Korrektheit des schlussfolgernden Arguments, die Beinhaltung eines empirischen Gehalts und eines allgemeinen Gesetzes des Explanans sowie die Korrektheit des Aussagegehalts des Explanans, vgl. *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 136; *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 123; *Hempel/Oppenheim*, Philo. Sci. 1948, S. 135, 137.

203 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 139 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Statistische Erklärungsmodelle sehen zwar niedrigere Anforderungen an den Wahrscheinlichkeitsgrad vor, dass das Explanan das Explanandum bedingt,[204] setzen aber dennoch eine Ursache-Wirkung-Beziehung voraus.

Wissenschaftstheoretisch erfordert eine Erklärung demzufolge, dass die Frage nach der Ursache beantwortet wird.[205] Die Untersuchung von Korrelationen kann – je nach Forschungsfrage – zwar auch Ziel einer wissenschaftlichen Analyse sein,[206] von größerem wissenschaftlichem Interesse sind aber die Kausalitäten, denn nur sie vermögen es, Erklärungen zu liefern. Die Wissenschaft strebt gerade danach, zu erforschen, worin ein Phänomen begründet liegt, anstatt sich mit dessen Existenz zu begnügen.[207] Das zeigt die hervorgehobene wissenschaftstheoretische Bedeutung der Erforschung von Kausalitäten.

Technisch ist der Nachweis von Kausalitäten jedoch äußerst anspruchsvoll. Er erfolgt meist experimentell mittels Kontrollgruppen oder Pfadanalysen[208] und erfordert eine sachlogische Interpretation von Korrelationen.[209] Auch die Rechtswissenschaft hat das Problem der schwierigen Nachweisbarkeit von Kausalitäten erkannt und sieht daher auf Definitions- wie auch auf Beweisebene Erleichterungen vor, indem sie nicht den natur-

---

204 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 141 f.

205 *Kornmesser/Büttemeyer,* Wissenschaftstheorie, S. 135; in Bezug auf die empirische Sozialforschung *Kühnel/Dingelstedt* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1401, 1401; *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 110; grundlegend *Hempel/Oppenheim*, Philo. Sci. 1948, S. 135, 135; anschaulich auch *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 110 ff.

206 *Mainzer* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 7. Genannt seien etwa sog. Warenkorbanalysen, die das Kaufverhalten analysieren. Dort genügt es zu wissen, dass bestimmte Produkte häufig gemeinsam erworben werden, sie also korrelieren. Anschaulich dazu *W. Zimmer,* Ansturm der Algorithmen, S. 259 f.

207 *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 110 f., 320 f.

208 Zur Informatik *Borgelt/Braune/Kruse* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 279, 320 ff.; zur Sozialforschung *Kühnel/Dingelstedt* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1401, 1407 ff.

209 *Weiber/Mühlhaus* in: Weiber/Mühlhaus (Hrsg.), Strukturgleichungsmodellierung, S. 9, 16.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

wissenschaftlichen Nachweis einer Ursache-Wirkungs-Beziehung erfordert, sondern lediglich eine an Sicherheit grenzende Wahrscheinlichkeit.[210]

2. Einordnung und Validität des Ergebnisses der TDM-Analyse

Mathematisch gesehen untersucht das TDM statistische Assoziationen, d. h. stochastische Abhängigkeiten, statt Kausalitäten.[211] Das gilt auch beim Einsatz von Methoden des maschinellen Lernens, diese können nach dem aktuellen Stand der Technik keine Kausalitäten nachweisen.[212] Es gibt bislang nur erste Ansätze, die gezielt Kausalitäten sichtbar machen wollen.[213]

Weil ein erhebliches Maß an Denkleistung notwendig ist, um eine Bedingung wegzudenken bzw. um zu überprüfen, ob die eine Bedingung auch ohne die andere denkbar wäre, stellt sich die Erforschung von Kausalitäten besonders schwierig dar.[214] Insofern stellt sich die Frage, ob die fehlende Fähigkeit von Systemen, die auf maschinellem Lernen basieren, Kausalitäten nachzuvollziehen, darauf zurückzuführen ist, dass Maschinen Texte oder Daten nicht bzw. anders verstehen als der Mensch:[215] Das

---

210 Verwiesen sei diesbezüglich auf die auch als Äquivalenztheorie bekannte *Conditio sine qua non-Formel*, so die rechtsgebietsübergreifende herrschende Auffassung in Literatur und Rechtsprechung, vgl. u. a. *Oetker* in MüKo, BGB, § 249 Rn. 103; *Freund* in MüKo, StGB, Vor § 13 Rn. 333 ff. Danach ist ein Ereignis kausal für das andere, wenn es nicht weggedacht werden könnte, ohne dass das andere mit an Sicherheit grenzender Wahrscheinlichkeit entfiele. Dieser Maßstab wird durch weitere Kriterien eingegrenzt, z. B. um unwahrscheinliche Kausalverläufe herauszufiltern, *Oetker* in MüKo, BGB, § 249 Rn. 109 ff.

211 *W. Zimmer,* Ansturm der Algorithmen, S. 205; *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 82

212 *Käde/Maltzan*, CR 2020, S. 66, 72 Rn. 33 ff.; *Ebers* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 3 Rn. 16 bezeichnet Big Data Analysen sogar als „*Primat der Korrelation über die Kausalitäten*", ebenso *Mainzer* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 6 f.: „*Datenkorrelationen können Hinweise auf kausale Zusammenhänge liefern, müssen es aber nicht*"; anhand von Beispielen aus der Datenwirtschaft erläutert das *Mayer-Schönberger/Cukier,* Big Data, S. 161 ff.

213 Z. B. von *Pearl* und *Mackenzie*, *Pearl/Mackenzie,* The book of why, S. 11 ff. sowie *Mainzer* in: Mainzer (Hrsg.), Philosophisches Handbuch Künstliche Intelligenz, S. 1, 7 f., vgl. auch *Peters/Janzing/Schölkopf,* Elements of causal inference sowie *Armbruster/Knop,* Wie intelligent werden Computer, Herr Schölkopf?, Minute 2.30.

214 Diese Bedingungen werden als „*kontrafaktuale*" *Bedingungen* bezeichnet, *Pearl/Mackenzie,* The book of why, S. 16.

215 *Käde/Maltzan*, CR 2020, S. 66, 71 Rn. 34.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Verstehen eines Textes erfolgt beim Menschen stets im Kontext des bisherigen Erfahrungsschatzes. Ein breiteres Wissen kann unter Umständen zum Erlangen von völlig anderen Erkenntnissen führen. Ähnlich verhält es sich mit dem maschinellen Lernen, denn je nach Umfang und Spezifität des Trainings versteht das jeweilige Modell den analysierten Text völlig unterschiedlich. Den Modellen fehlt jedoch gerade das Wissen um die mögliche Beschränktheit des eigenen Erfahrungsschatzes.[216] Ob das alte Verständnis von „*Verstehen*" und „*Intelligenz*" allerdings in Zeiten, in denen KI-Systeme Zeitungsartikel verfassen oder berühmte Werke im Stil des jeweiligen Autors fortführen können,[217] aufrechterhalten werden kann, ist anzuzweifeln, denn die Andersartigkeit des Verständnisses desselben Textes des Algorithmus im Vergleich zum Menschen beinhaltet nicht zwangsläufig ein geringeres Potenzial. Das „*Verstehen*" durch Maschinen steht dem Verstehen durch den Menschen daher nicht in jedem Fall nach.

Im Rahmen von TDM-Analysen stellt die Ergebnisanalyse einen erheblichen Anteil des Forschungsprozesses dar. Das ist darauf zurückzuführen, dass durch die TDM-Analyse eine große Anzahl an Zusammenhängen hervorgebracht wird, die erst in Bezug zur eigentlichen Forschungsfrage gesetzt und auf ihre wissenschaftliche Relevanz überprüft werden müssen.[218] Auf Grundlage der durch die Analyse erlangten Korrelationen ist

---

216 Das kann sich möglicherweise ändern, wenn zu einem späteren Zeitpunkt eine starke KI technisch realisierbar ist, wobei sich die Forschung uneins ist, ob auch eine starke KI ein Bewusstsein entwickeln kann, was wohl für ein „*Verständnis*" von Inhalten, wie es der Mensch entwickelt, erforderlich ist, vgl. *Niederée/Neidl* in Ebers/Heinze/Krügel, Künstliche Intelligenz und Robotik § 2 Rn. 2. Von führenden Forschern wird das (Nicht-)Verstehen von Kausalzusammenhängen als die entscheidende Hürde betrachtet, die überwunden werden muss, um Maschinen menschliche Intelligenz zu verleihen. *Pearl/Mackenzie*, The book of why, S. 10 f. bezeichnen eine fehlende Fähigkeit von Machine-Learning-Systemen, Kausalitäten zu verstehen, sogar als die größte Barriere zum Erreichen menschlicher Intelligenz. Das Vorstellungsvermögen des Menschen stellt ihnen zufolge in Bezug auf die sog. kontrafaktualen Konditionen gerade die kognitiven Vorsprünge des Menschen dar, die es erst ermöglicht, überhaupt zu forschen und die den Menschen klüger als seine Daten macht, *Pearl/Mackenzie*, The book of why, S. 16, 21.

217 Vgl. *Lobe*, Frankfurter Allgemeine Sonntagszeitung 20.09.2020, S. 36 zum Textgenerator GPT-3.

218 *Abts/Mülder*, Grundkurs Wirtschaftsinformatik, S. 286: „*Allgemein muss mit einer hohen Rate an scheinbar trivialen Zusammenhängen, d.h. Sachverhalten, die bereits bekannt oder offensichtlich sind, gerechnet werden. Letztlich obliegt es dem erfahrenen Anwender, nichttriviale Zusammenhänge als solche zu erkennen, um dann die wirklich wertvollen Teilergebnisse näher zu untersuchen*".

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

es jedenfalls möglich, auf Kausalitäten zu schließen. Darauf aufbauend ist die durch das TDM erlangte Erkenntnis nicht nur bei der Anwendung von statistischen Methoden,[219] sondern auch beim Einsatz von Methoden des maschinellen Lernens nicht unmittelbar im Analyseergebnis, sondern vielmehr in der vollendeten Interpretation der Ergebnisse zu sehen.[220] Das senkt aber nicht den Wert der mittels TDM aufgefundenen Korrelationen: Unter der Prämisse, dass es im wissenschaftlichen Prozess ein normaler Vorgang ist, Kausalitäten separat zu erforschen, wäre ausschlaggebendes Kriterium für die Einordnung des Potenzials von TDM nicht die Frage, ob sie – im Gegensatz zu anderen Technologien – Kausalitäten hervorbringen kann, sondern vielmehr die Frage, ob die Korrelationen dem Menschen tatsächlich nicht zugänglich sind. Insofern sollten Kausalitätsnachweise auch bei TDM-Prozessen als Aufgabe der Interpretation angesehen werden.[221]

Für die Beantwortung der Frage, ob die im TDM gewonnenen Erkenntnisse dem Menschen nicht zugänglich sind, ist in erster Linie die Größe der Datengrundlage entscheidend. Je größer die Datengrundlage ist, desto eher ist eine Unzugänglichkeit der Erkenntnisse für den Menschen anzunehmen. Eine Zugänglichkeit der Erkenntnisse für den Menschen kann also nicht generell ausgeschlossen werden,[222] eine Unzugänglichkeit ist aber dennoch – jedenfalls aus praktischen Gründen – meistens die Realität. Darin ist das besondere Potenzial zu sehen. Die neue Erkenntnis kann daneben auch im neuronalen Netz selbst liegen, welches die Abstrahierungen in den Gewichten speichert.[223] Diese nutzbar zu machen ist aber angesichts der tiefen neuronalen Netze ungemein schwierig und dürfte wenig Antworten auf konkrete Fragen liefern.

Eine besondere Schwierigkeit, wenn für das TDM Algorithmen genutzt werden, die auf maschinellem Lernen basieren, liegt in der als solche be-

219 Ein Beispiel dafür ist die beschriebene Kookkurenzanalyse, die hervorbringt, ob linguistische Tokens häufig gemeinsam auftreten, bei der aber erst im Anschluss separat geprüft werden muss, ob auch ein semantischer Zusammenhang besteht, der über eine bloße parallele Häufung hinausgeht, *Manderscheid* in: Baur/Blasius (Hrsg.), Handbuch Methoden der empirischen Sozialforschung, S. 1103, 1110 f.

220 So wohl auch *Bramer,* Principles of Data Mining, S. 2, Figure 1.1.

221 *Stegmüller,* Probleme und Resultate der Wissenschaftstheorie und analytischen Philosophie, S. 111.

222 *Giudici,* Applied data mining, S. 1.

223 *Stiemerling*, CR 2015, S. 762, 764.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

zeichneten „*Black-Box*“,[224] d. h. in der schwierigen Nachvollziehbarkeit.[225] Besonders schwer fällt der Nachvollzug, wenn tiefe künstliche neuronale Netze eingesetzt werden.[226] Das gilt insbesondere bei Analysen des unüberwachten Lernens, bei dem keine Kategorisierung durch den Menschen vorgegeben ist: Im Gegensatz zum überwachten Lernen wurde beim unüberwachten Lernen nicht bereits an die Relevanz bestimmter Kriterien gedacht – dementsprechend fällt das Ziehen von Rückschlüssen ohne Kategorisierung schwerer.[227] Die schwierige Erklärbarkeit liegt demnach gerade in denjenigen Mechanismen, die sich der Mensch bei umfangreichen Datenmengen zunutze macht, die er manuell nicht bewältigen könnte. Um dem Problem der fehlenden Nachvollziehbarkeit zu begegnen, wird an Erklärungs-Algorithmen geforscht, die z. B. relevante Stellen mit den erkannten Kategorien markieren, Schichten visualisieren oder die den Einfluss der Gewichte auf die Entscheidungen veranschaulichen.[228] Spezialisten können die Prozesse theoretisch bereits jetzt nachvollziehen.[229]

---

224 *Zech*, ZfPW 2019, S. 198, 202. Die Auffassung, es handele sich im Bereich des maschinellen Lernens grundsätzlich und in jedem Fall um eine Black-Box, erfährt starke Kritik, vgl. *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 59; *Käde/Maltzan*, CR 2020, S. 66, 67 Rn. 7. Demgegenüber werden dennoch Stimmen lauter, die – schon aus Gründen der Akzeptanz, die zur Gesellschaftsfähigkeit von KI beitragen soll, im Wissenschaftsbereich aber nicht infrage steht – eine „*Explainable AI*“ oder „*XAI*“ fordern, dazu *Käde/Maltzan*, CR 2020, S. 66, 67 f. Rn. 5, 12 ff. Auch die Europäische Kommission fordert in ihrem Weißbuch zur künstlichen Intelligenz im Rahmen der EU-Digitalstrategie Transparenz und Vertrauenswürdigkeit, Europäische Kommission, Weißbuch zur Künstlichen Intelligenz - COM(2020) 65 final, S. 1 ff.

225 *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 59.

226 Im Falle von Backpropagation Netzwerken, d. h. künstlichen neuronalen Netzen, die mit einer Rückkopplung arbeiten und deren Gewichte und Entscheidungen sich fortlaufend anpassen, ist die Nachvollziehbarkeit zusätzlich erschwert, *Käde/Maltzan*, CR 2020, S. 66 ff., 66, 69 Rn. 7, 21 f. Auch bei Entscheidungsbäumen, einer weiteren Kategorie von Algorithmen aus dem maschinellen Lernen, ist das u. a. der Fall, wenn der Entscheidungsbaum besonders viele Ebenen hat oder Zufallsvariablen beinhaltet (sog. *Random forests*), *Käde/Maltzan*, CR 2020, S. 66, 67 Rn. 7.

227 *Stiemerling*, CR 2015, S. 762, 764; *Stiemerling* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.1 Rn. 58 ff.; *Körner* in Kaulartz/Braegelmann, Rechtshandbuch AI, Kapitel 2.4 Rn. 23.

228 *Käde/Maltzan*, CR 2020, S. 66, 70 Rn. 25 ff.

229 *Käde/Maltzan*, CR 2020, S. 66, 67 Rn. 7. Das muss den Anforderungen an den wissenschaftlichen Nachweis genügen, schließlich ist es nicht Anspruch an die Wissenschaft, ihre Erkenntnisse für jedermann verständlich zu gestalten. Es genügt, wenn Experten dazu in der Lage sind: Die entsprechenden Peer-Review-

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

E. Weitere grundlegende Begriffsverständnisse

Zum Abschluss der technischen Grundlagen der Untersuchung sind noch einige weitere Verständnisse von Begriffen zu erläutern.

I. Begriffsverständnis von Algorithmus und Computerprogramm

Die TDM-Analyse erfolgt algorithmusbasiert. Ein Algorithmus ist eine mathematische Formel, d. h. eine Vorgabe bzw. ein Schema eines Rechenvorgangs.[230] Im Gegensatz dazu ist ein Computerprogramm eine mögliche Umsetzung des Algorithmus, sodass der Computer nach diesem agieren kann, der Algorithmus wird also in ein Computerprogramm implementiert.[231]

II. Begriffsverständnis von Information, Wissen und Daten

Die These, das TDM extrahiere Wissen, erfordert eine Differenzierung, wie die Begriffe der Daten, der Informationen und des Wissens zueinander stehen. Zu dieser Abgrenzung existiert eine Fülle von Ansichten, die in dieser Bandbreite für die vorliegende Untersuchung nicht zielführend darzustellen sind. Relevant ist vielmehr eine knappe und kontextbezogene Einordnung.

Aus informationstechnischer Perspektive sind Daten maschinenlesbare, d. h. in Binärcode vorliegende Informationen, z. B. Werte, Messdaten,

---

Verfahren werden von diesen Experten der jeweiligen Fachgebiete übernommen. Diese sind dazu in der Lage, zu verifizieren, ob die Analysen ordnungsgemäß abgelaufen sind, die Forschungsdaten hinreichend vorverarbeitet waren, geeignete Analyse-Tools verwendet wurden und Zufallshäufungen ausgeschlossen wurden. Insofern stellt eine vermeintlich fehlende Nachvollziehbarkeit im Wissenschaftsbereich theoretisch nicht die Validität der wissenschaftlichen Erkenntnis infrage, vgl. dazu später auf S. 268 ff.

230 So der Duden, Definition des Algorithmus, abrufbar unter: https://www.duden.de/rechtschreibung/Algorithmus; *Gumm/Sommer et al.*, Einführung in die Informatik, S. 91; *Ernst/Schmidt/Beneken*, Grundkurs Informatik, S. 6.

231 *Ernst/Schmidt/Beneken*, Grundkurs Informatik, S. 6; Käde, Kreative Maschinen und Urheberrecht, S. 43; *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 69a Rn. 22; *Käde*, Kreative Maschinen und Urheberrecht, S. 43; *Ernst/Schmidt/Beneken*, Grundkurs Informatik, S. 6.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Texte oder Bilder.[232] Unter Informationen ist dagegen gerade ein in Daten verkörperter oder jedenfalls verkörperungsfähiger Inhalt zu verstehen.[233] Der Begriff des Wissens, der auf *Platon* zurückgeht, der ein Verständnis von Wissen als „*wahre, begründete Überzeugung*" vorschlug,[234] ist vielschichtig und kann sich auf die Erkenntnis als solche, aber auch auf das Wissen als geistigen Prozess beziehen.[235] Festzuhalten ist jedoch, dass Wissen von der wissenschaftlichen Prüfung bzw. Interpretation[236] vorhandener Information abhängt. Angeknüpft wird beim Begriff des Wissens also an die subjektive Rezeption bzw. vernetzte Wahrnehmung und Einordnung durch den Menschen.[237] Wissen entsteht also nicht (immer) allein durch eine konkrete Informationsaufnahme, sondern durch einen Prozess, bei dem sich aus der Aufnahme einer großen Bandbreite einzelner Informationen und deren kognitiver Verarbeitung bzw. aus der Schlussfolgerung ein breiteres Wissen bildet,[238] der in den wissensbasierten Systemen bzw. Expertensystemen gerade repliziert werden soll.

---

232 *Zech,* Information als Schutzgegenstand, S. 32 f., 55; *Wiegerling/Nerukar/Wadephul* in: Kolany-Raiser/Heil/Orwat/Hoeren (Hrsg.), Big Data, S. 401, 405.

233 *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 10 f.; *Martini/Kolain et al.*, MMR-Beilage 2021, S. 3, 3. *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 10 f.; *Martini/Kolain et al.*, MMR-Beilage 2021, S. 3, 3. Bei Informationen sind grundsätzlich drei Ebenen zu unterscheiden, namentlich die syntaktische, die semantische und die pragmatische Ebene: Die syntaktische Ebene betrachtet dabei nur das Gefüge der Zeichen, d. h. die formale Richtigkeit, die semantische Ebene bezieht sich auf die inhaltliche Bedeutung und die pragmatische Ebene auf die Wahrnehmung und Verarbeitung durch den Rezipienten, also durch den Menschen, *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 10 f.

234 *Zech,* Information als Schutzgegenstand, S. 28; *Gottschalk-Mazouz* in: Ammon (Hrsg.), Wissen in Bewegung, S. 21, 22 ff.; vgl. dazu *Platon,* Theätet, in der sich Platon in einem Dialog zwischen Sokrates und einem Mathematiker und dessen Schüler über die Unterscheidung von Wissen, Wahrnehmung und Meinung einem Verständnis von Wissen annähert.

235 *Schnädelbach,* Erkenntnistheorie zur Einführung, S. 32 f.; *Peter Baumann,* Erkenntnistheorie, S. 33 ff.; *G. Ernst,* Einführung in die Erkenntnistheorie, S. 55 ff. Die wissenschaftstheoretische Diskussion, wie sich Wissen und Wahrheit zueinander verhalten, kann an dieser Stelle nicht geführt werden. Vgl. dazu jedoch die in dieser Fn. genannten Quellen.

236 *Zech,* Information als Schutzgegenstand, S. 28; *Wiegerling/Nerukar/Wadephul* in: Kolany-Raiser/Heil/Orwat/Hoeren (Hrsg.), Big Data, S. 401, 414; *Bramer,* Principles of Data Mining, S. 3.

237 *Abts/Mülder,* Grundkurs Wirtschaftsinformatik, S. 12.

238 Diesbezüglich wird zwischen implizitem und explizitem sowie prozeduralem Wissen unterschieden, *Nebel/Wölfl* in: Görz/Schmid/Braun (Hrsg.), Handbuch der Künstlichen Intelligenz, S. 27, 106 f., 110 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Die Unterscheidung von Daten, Wissen und Information wird in Gesellschaft und Wissenschaft, aber gerade auch in den datenbezogenen Rechtsgebieten, nicht einheitlich vorgenommen.[239] Im Urheberrecht treten Daten als solche selten in Erscheinung, relevant werden sie ausdrücklich nur im Bereich des Datenbankherstellerschutzes, § 87a Abs. 1 S. 1 UrhG. Die zugrundeliegende Datenbank-RL definiert dabei nicht, was sie unter Daten versteht, scheint aber davon auszugehen, dass es bei Daten nicht auf ihren Inhalt ankommt, so stellt sie klar, dass die Datenbank-RL keine „*Ausdehnung des urheberrechtlichen Schutzes auf reine Fakten oder Daten*" beinhaltet.[240] In der europäischen Datenstrategie[241] werden Daten schlicht nicht definiert. Gerade im Datenschutzrecht ist trotz Uneinigkeit über den Informations- bzw. den Datenbegriff jedoch kennzeichnend, dass Informationen stets ein Zuordnungsobjekt oder -subjekt haben,[242] wobei es nur auf das Vorhandensein eines Aussagegehalts der Informationen und nicht seiner Wahrheit oder Erwiesenheit ankommt.[243] Herausarbeiten lässt sich also, dass selbst in Kontexten, in denen die Begriffe der Information und des Datums synonym verwendet werden, die Tendenz erkennbar ist, dass Informationen gemeint sind, wenn der Gehalt, der vermittelt wird, thematisiert wird.

---

239 *Wirtz,* Die Kommerzialisierung kultureller Informationen der öffentlichen Hand, S. 29 f.; *Ingold,* Desinformationsrecht, S. 21 ff. Das gilt z. B. für das Datenschutzrecht, *Spiecker gen. Döhmann*, RW 2010, S. 247, 250 f.; *Determann*, MMR 2018, S. 277, 277; *Determann*, ZD 2018, S. 503, 504; *M. Karg* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art. 4 Nr. 1 DSGVO Rn. 25; *Klar/Kühling* in Kühling/Buchner, DSGVO/BDSG, Art. 4 Abs. 1 Rn. 8; EuGH, ECLI:EU:C:2017:994 – *Nowak* = NJW 2018, S. 767, 768 f. Rn. 34; für das Informationsrecht *Schoch* in: Vereinigung der Deutschen Staatsrechtslehrer (Hrsg.), Der Grundrechtseingriff. Öffentlich-rechtliche Rahmenbedingungen einer Informationsordnung, S. 158, 160 f., 166 f.; *Schoch* in Schoch, IFG, § 2 Rn. 17; *Ingold,* Desinformationsrecht, S. 21. Auch im Bereich der grundrechtlichen Informationsfreiheit, Art. 5 Abs. 1 S. 2 GG, gibt es keine einheitliche Definition der Information, *Grabenwarter* in Dürig/Herzog/Scholz, GG, Art. 5 Abs. 1, 2 Rn. 999. f.

240 Richtlinie (EU) 96/6/EG, Datenbank-RL Erw. 45.

241 Europäische Kommission, Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen - Eine europäisch Datenstrategie, COM(2020) 66 final (19.02.2020).

242 Im Datenschutzrecht sagen sie etwas über jemanden, also eine Person, aus.

243 *M. Karg* in Simitis/Hornung/Spiecker gen. Döhmann, Datenschutzrecht, Art. 4 Nr. 1 DSGVO Rn. 29.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Resultierend aus diesen Erkenntnissen sollen die Begriffe in der folgenden Untersuchung kontextbezogen verwendet werden, sodass Daten dann adressiert werden, wenn maschinenlesbare Informationen ungeachtet ihres Inhalts relevant sind. Von Information kann hingegen erst dann gesprochen werden, wenn Daten einen Inhalt vermitteln oder sich auf bestimmte Zustände beziehen. Das gilt erst recht für Wissen, bei dem eine subjektive Rezeption und eine Einordnung der Information als Wissen notwendig sind.

III. Begriffsverständnis von Forschungsdaten

Der Begriff der „*Forschungsdaten*" ist ebenso wie die bereits thematisierten Begriffsverständnisse konturenlos. Im Sinne einer möglichst großen Bandbreite an Fallkonstellationen und in Anlehnung an die Vielfältigkeit an Art und Ursprung sowie Zustand der betreffenden Daten soll der Arbeit im Folgenden ein möglichst weites Verständnis von Forschungsdaten zugrunde gelegt werden.[244] Aus dem Begriff selbst ist jedenfalls zu folgern, dass Forschungsdaten maschinenlesbare sowie körperlich fixierte Informationen sind, die zu wissenschaftlichen Zwecken gesammelt, aufbereitet oder verwendet werden sollen oder es bereits wurden. Insofern sind sowohl Rohdaten (Primärdaten) als auch verarbeitete Daten (Sekundärdaten) darunter zu fassen.[245] Insofern kann der Begriff eine Vielzahl unterschiedlicher urheberrechtlicher Schutzgegenstände oder auch gemeinfreier Ressourcen meinen.[246]

---

244 Diesen Ansatz verfolgt auch Kuschel, die unter Forschungsdaten alle Daten versteht, die „*Gegenstand eines wissenschaftlichen Erkenntnisprozesses sind*" *Kuschel*, OdW 2020, S. 43, 43.

245 *Paul Baumann/Krahn/Lauber-Rönsberg,* Forschungsdatenmanagement und Recht: Datenschutz-, Urheber- und Vertragsrecht, S. 16; *Brase* in: Neuroth/Oßwald/Scheffel/Strathmann/Jehn (Hrsg.), Eine kleine Enzyklopädie der digitalen Langzeitarchivierung, S. 57, 59 f.

246 Diese Auffassung vertritt auch der Bundesrat: „*Die Sammlung und Aufbereitung von Forschungsdaten, soweit sie urheberrechtlich geschützt sind,*", BR-Drs. 142/21 (Beschluss), S. 5. Ein engeres Verständnis hat hingegen *T. Hartmann*, InTeR 2013, S. 199, 200, der sich hauptsächlich auf Messdaten oder sonstige numerischen Werte bzw. Ergebnisse aus Versuchen bezieht.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

F. Zwischenfazit

TDM ist als Technologie äußerst vielfältig und beinhaltet die Trainingsprozesse, die zum maschinellen Lernen zählen bis hin zu geisteswissenschaftlichen Analysen. Wie andere Forschungsmethoden erfordern auch die mit TDM erlangten Analyseergebnisse eine dezidierte Interpretation und Prüfung. Das besondere Potenzial von TDM-Analysen liegt demnach nicht darin, dass sie im Gegensatz zu anderen Methoden Kausalzusammenhänge hervorbringen könnten, sondern darin, dass sie enorme Daten- und Textmengen bewältigen können. Die Probleme, die bei dem Nachvollzug der Analysen auftreten, liegen demnach in der Tatsache begründet, dass die Methoden gerade dazu eingesetzt werden, wozu der Mensch nicht oder nicht unter angemessenem Aufwand fähig ist. Dennoch ist der Prozess des TDMs kein feststehender Begriff und auch – insbesondere in den digitalen Geisteswissenschaften als Forschungsfeld, das sich selbst noch in einer frühen (Entwicklungs-)Phase befindet – keine feststehende Methode. Prägend ist, dass die Technologien des TDMs und der betroffenen Teilbereiche des Forschungsbereichs der KI sich nicht nur begrifflich überschneiden, sondern sich auch technisch gegenseitig positiv beeinflussen. Folglich hat der Stand der KI-Forschung eine große Bedeutung für das TDM, auch wenn TDM-Analysen ohne die Nutzung von Methoden des maschinellen Lernens vorgenommen werden können.

§ 4 Ein Urheberrecht für die Wissenschaft?

Das TDM wird in dieser Untersuchung sowohl aus seiner wissenschaftlichen als auch aus seiner innovationspolitischen Seite beleuchtet. Um das TDM als wissenschaftliche Methode untersuchen zu können, ist notwendig zu verstehen, welche besonderen und schutzwürdigen Interessen das Urheberrecht in Bezug auf diesen speziellen Bereich urheberrechtlichen Werkschaffens, aber auch des Werknutzens berücksichtigen muss und wie sich die systematischen Grundlagen auf die Steuerung der Wissenschaft durch das Urheberrecht auswirken. Zu beleuchten ist daneben, wie und wieso sich die urheberrechtliche Systematik in Bezug auf die Wissenschaft vom sonst tendenziell urheberzentrierten Recht unterscheidet. Besondere Wertungen kann eine Betrachtung auf grundrechtlicher Ebene und auf Ebene des einfachen Urheberrechts inklusive der unionsrechtlichen Einflüsse ergeben.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

A. Text- und Data-Mining zur wissenschaftlichen Forschung

I. Das Begriffsverständnis von Wissenschaft im Grundgesetz

Das Grundgesetz, das in Art. 5 Abs. 3 S. 1 Alt. 2 Wissenschaft, Forschung und Lehre schützt, versteht unter der Wissenschaft den *„ernsthaften und planmäßigne Versuch zur Ermittlung der Wahrheit"*[247] und unter der Forschung eine *„geistige Tätigkeit, die das Ziel verfolgt, auf wissenschaftliche Art und Weise neue Erkenntnisse zu gewinnen"*.[248] Auf die Wissenschaftsfreiheit, Art. 5 Abs. 3 S. 1 Alt. 2 GG kann sich dabei grundsätzlich jeder berufen, der wissenschaftlich arbeitet.[249] In jedem Fall muss die wissenschaftliche Erkenntnisgewinnung eine gewisse Verlässlichkeit aufweisen und reproduzierbar sein.[250] Kennzeichen ist die *„prinzipielle Unabgeschlossenheit"* der wissenschaftlichen Erkenntnis.[251] Die Wissenschaftsfreiheit schützt die *„wissenschaftlichen Eigengesetzlichkeiten"*,[252] d. h. diejenigen Prozesse, die der wissenschaftlichen Betätigung eigentümlich sind. Dazu zählt auch eine *„nachvollziehbare Gesetzmäßigkeit"*.[253] Ein wichtiges Kennzeichen ist dabei die Anerkennung einer Forschung in der Wissenschaftsgemeinschaft bzw. der *„scientific community"*.[254] Damit definiert die Wissenschafts-

---

247 BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176; BVerfG, Beschl. v. 1. 3. 1978 - 1 BvR 174, 178, 191/71; 333/7 = NJW 1978, S. 1621, 1621; *Gärditz* in Dürig/Herzog/Scholz, GG, Art. 5 Abs. 3 Rn. 55. Das Verständnis geht ursprünglich auf *Smend* zurück und bezog sich noch auf die Weimarer Reichsverfassung, *Schulz-Prießnitz,* Einheit von Forschung und Lehre, S. 30.

248 *Gärditz* in Dürig/Herzog/Scholz, GG, Art. 5 Abs. 3 Rn. 94.

249 BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176; BVerfG, Beschl. v. 16.1.1963 - 1 BvR 316/60 = NJW 1963, S. 899, 899.

250 *Gärditz* in Dürig/Herzog/Scholz, GG, Art. 5 Rn. 76; BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176.

251 BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176.

252 *Gärditz* in Dürig/Herzog/Scholz, GG, Art. 5 Abs. 3 Rn. 57; BVerfG, Beschl. v. 16.1.1963 - 1 BvR 316/60 = NJW 1963, S. 899, 1176.

253 *Kempen* in Epping/Hillgruber, GG, Art. 5 Rn. 181; *Pernice* in Dreier, GG, 2. Aufl. 2004, Art. 5 III (Wissenschaft) Rn. 27.

254 *Gröschner* in GS Helm, S. 23 ff.; *Kirchhof,* Wissenschaft in verfaßter Freiheit, S. 2 betont ebenso die Bedeutung der *„eigenen Gesetzmäßigkeit; Ruffert* in: Vereinigung der Deutschen Staatsrechtslehrer (Hrsg.), Kultur und Wissenschaft, S. 146, 154; *Trute,* Die Forschung zwischen grundrechtlicher Freiheit und staatlicher Institutionalisierung, S. 106, weiter zur Subkultur der Wissenschaftsgemeinschaft bzw. scientific community auf S. 64 ff.; *Schmidt-Aßmann* in FS Meusel, S. 217, 219 f., der betont, dass Wissenschaft erst durch die Kommunikation zur solchen wird, relevant sei, *„ob und inwieweit ein Forschungsergebnis in der*

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

gemeinschaft selbst, was Wissenschaft ist. Das hat zur Folge, dass der Begriff der Wissenschaft im Rahmen der Wissenschaftsfreiheit ähnlich unbestimmt ist wie der Begriff der Kunst im Rahmen der Kunstfreiheit (Art. 5 Abs. 1 S. 1 Alt. 1 GG).

Das dargestellte Verständnis wissenschaftlicher Tätigkeit, d. h. die tatsächliche Erkenntnissuche, bezieht sich vorwiegend auf ihren Unterfall der Forschungsfreiheit.[255] Dabei wird nicht nur die unmittelbare erkenntnisbringende Tätigkeit durch die Wissenschaftsfreiheit bzw. die Forschungsfreiheit privilegiert, sondern auch Vorbereitungshandlungen inklusive der Materialsammlung.[256] Auch kommerzielle Beweggründe schließen den Schutz durch die Wissenschaftsfreiheit nicht aus, auch wenn die (kommerzielle) Verwertung der Forschung nicht durch die Wissenschaftsfreiheit, sondern durch die Berufsfreiheit, Art. 12 Abs. 1 GG, geschützt wird.[257]

Wissenschaft real als solches diskutiert wird", es komme also gerade auf die Vermittelbarkeit in der scientific community an; auch *Ridder,* Die soziale Ordnung des Grundgesetzes, S. 175 betont die Relevanz des wissenschaftlichen Selbstverständnisses als schutzbereichsdefinierend: „*Alles, was in diesem Prozess als »wissenschaftlich« akzeptiert wird, ist von den außerwissenschaftlichen Instanzen als wissenschaftlich zu akzeptieren. Der Wissenschaftsbegriff steht also nur der Wissenschaft zur Verfügung. Wissenschaftsfreiheit impliziert mithin das außerwissenschaftliche Definitionsverbot über Wissenschaft*"; gegen eine Berücksichtigung kommunikativer Elemente Schulz-Prießnitz, Einheit von Forschung und Lehre, S. 32 f. für eine Indizwirkung Fehling in BK GG, Art. 5 Abs. 3 (Wissenschaftsfreiheit) Rn. 61 f., der daneben auch „*Schutzzweck und grundrechtstypische Gefährdungslagen*" einbezieht.

255 *Von Coelln* in Berliner Kommentar, GG, Art. 5 Rn. 28 f.; *Jarass* in Jarass/Pieroth, GG, Art. 5 Rn. 136; *Wendt* in v. Münch/Kunig, GG, 7. Aufl. 2021, Art. 5 Rn. 156, 159; *R. Dreier*, DVBl 1980, S. 471, 471; BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176 f.; BVerfG, Beschl. v. 20.10.1982 – 1 BvR 1467/80 = BeckRS 1982, 106553; BVerfG, Beschl. v. 29.06.1983 - 2 BvR 720/79, 2 BvR 725/79, 2 BvR 742/79,2 BvR 1579/79, 2 BvR 1582/79, 2 BvR 826/80, 2 BvR 1168/80 = NJW 1984, S. 912, 913 ff.; BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1176.

256 *Chr. Starck/A. L. Paulus* in v. Mangoldt/Klein/Stark, GG, Art. 5 Rn. 474; *Jarass* in Jarass/Kment/Pieroth, Grundgesetz für die Bundesrepublik Deutschland Art. 5 Rn. 138.

257 Zu dieser Fragestellung existiert eine Fülle an Meinungen, aus der sich die o. g. These jedoch überwiegend herauskristallisiert, vgl. *Bäuerle* in: Britz (Hrsg.), Forschung in Freiheit und Risiko, S. 1, 10; *Blankenagel*, AöR 2000, S. 70, 94; *Gärditz*, ZBR 2009, S. 145, 147; *Gärditz*, JZ 2007, S. 521, 522; *Kleindiek,* Wissenschaft und Freiheit in der Risikogesellschaft, S. 320; *Ruffert* in: Vereinigung der Deutschen Staatsrechtslehrer (Hrsg.), Kultur und Wissenschaft, S. 146, 158 f.; *Trute,* Die Forschung zwischen grundrechtlicher Freiheit und staatlicher Institutionalisierung, S. 106; *Fehling* in BK GG, Art. 5 Abs. 3 Rn. 81; *Gärditz* in Dü-

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Insofern ist auch die sog. „*Zweck- oder Auftragsforschung*" erfasst, solange nach wissenschaftlichen Grundsätzen vorgegangen wird und nicht bloß bestehendes Wissen angewandt wird.[258]

II. Das Verständnis von Wissenschaft im Urheberrecht

Dem UrhG, in dem die Wissenschaft sowohl auf Ebene des Werkschutzes – insbesondere in § 2 UrhG – als auch in den Schrankenregelungen, insbesondere in den §§ 51, 60c, 60d UrhG, adressiert ist, liegt kein einheitliches Verständnis von Wissenschaft zugrunde.

Gerade Darstellungen wissenschaftlicher oder technischer Art, die in § 2 Abs. 1 Nr. 7 UrhG normiert sind, dienen viel eher der Informationsvermittlung und einer Belehrung, entscheidend ist insofern nicht die Wissenschaftlichkeit des dargestellten Inhalts, sondern die belehrende „*Darstellung in einer individuellen Form*".[259] Maßstab ist also gerade nicht der dargestellte Inhalt.[260] Im Gegensatz zum grundrechtlichen Verständnis ist also auf Schutzbereichsebene nicht auf den wissenschaftlichen Erkennt-

---

rig/Herzog/Scholz, GG, Art. 5 Abs. 3 Rn. 90. Ein noch weiteres Verständnis der Wissenschaftsfreiheit in kommerziellen Kontexten hat *Frieling*, GRUR 1987, S. 407, 408, der auch die kommerzielle Verwertung unter die Wissenschaftsfreiheit fasst.

258 So u. a. *Fehling* in BK GG, Art. 5 Abs. 3 Rn. 81; *Gärditz* in Dürig/Herzog/Scholz, GG, Art. 5. Rn. 101; *Chr. Starck/A. L. Paulus* in v. Mangoldt/Klein/Stark, GG, Art. 5 Rn. 477; *Wendt* in v. Münch/Kunig, GG, 7. Aufl. 2021, Art. 5 Rn. 156; *von Coelln* in Berliner Kommentar, GG, Art. 5 Rn. 32; *Dickert*, Naturwissenschaften und Forschungsfreiheit, S. 259 ff. spricht sich deswegen dafür aus, weil sich ein Wissenschaftsverständnis seit der Entstehung des Grundgesetzes gewandelt habe und diese Eigengesetzlichkeiten gerade den Schutzbereich definieren; *Kamp*, Forschungsfreiheit und Kommerz, S. 52 ff.; enger sieht das *Dähne*, Forschung zwischen Wissenschaftsfreiheit und Wirtschaftsfreiheit, S. 393 ff., der nur zweckfreie Forschung erfasst sieht.

259 *Bullinger* in Wandtke/Bullinger, UrhG, 5. Aufl. 2019, § 2 Rn. 132.

260 *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 226; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 223; BGH, Urt. v, 10.10.1991 - I ZR 147/89 – *Bedienungsanweisung* = GRUR 1993, S. 34, S. 35; BGH, Urt. v. 28.05.1998 - I ZR 81/96 – *Stadtplanwerk* = GRUR 1998, S. 916, 917; BGH, Urt. v. 15. 12. 1978 - I ZR 26/77 – *Flughafenpläne* = NJW 1979, S. 1548, 465; KG, Urt. v. 11.7.2000 - 5 U 3777/99 – *Memokartei* = GRUR-RR 2002, S. 91, 92; OLG München, Urt. v. 19.09.1991 - 6 U 2093/38 – *Rätsel* = GRUR 1992, S. 510, 510.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

nisprozess, sondern auf den belehrenden Charakter des Erzeugnisses abgestellt.[261]

Die urheberrechtlichen Schrankenbestimmungen zeigen hingegen stärkere Überschneidungen mit dem verfassungsrechtlichen Wissenschaftsbegriff. Das wird darauf zurückgeführt, dass sie einer grundrechtlichen Legitimierung bedürfen.[262] Insbesondere im Rahmen des wissenschaftlichen Großzitats, § 51 S. 2 Nr. 1 UrhG, werden Parallelen zum grundrechtlichen Wissenschaftsverständnis gezogen.[263] Dasselbe gilt in Bezug auf das Verhältnis von Wissenschaft zu Lehre und Forschung,[264] denn auch dort wird zwischen Wissenschaftsvermittlung (§ 60a UrhG) und wissenschaftlicher Forschung (§ 60c UrhG) differenziert. Unter der Wissenschaftlichkeit i. S. d. § 60c UrhG wird eine „*Tätigkeit*" verstanden, die „*aufgrund einer methodisch systematischen Vorgehensweise Erkenntnis anstrebt und auch vermittelt*".[265] Die Forschung hat dabei das Ziel der „*Wissensschaftsproduktion*" und der „*Gewinnung neuer Erkenntnisse*".[266]

Das Verständnis von der Wissenschaft, das den Schrankenbestimmungen zugrunde liegt, geht auf die europäischen Richtlinien, insbesondere die InfoSoc-RL, aber auch die neuere DSM-RL zurück, die ihrerseits im Lichte des unionsrechtlichen Verständnisses der Wissenschaftsfreiheit auszulegen sind, Art. 13 S. 1 GRCh.[267] Zentral ist danach das methodische und systematische Streben zur Gewinnung neuer Erkenntnisse in nachprüfbarer Art und Weise.[268] Die Differenzierung der deutschen Schran-

---

261 *Rieger,* Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 43. Es erweist sich insbesondere in der gerichtlichen Auslegungspraxis, dass das Verständnis wissenschaftlicher Werke im Sinne des UrhGs „*mit methodengerechter Wahrheitssuche und intrinsischer Wahrheitsliebe*" nicht so eng verbunden ist, wie es den Anschein erweckt, *Peukert*, JIPITEC 2013, S. 142, 144 Rn. 15.

262 *Rieger,* Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 42. Zum Verhältnis von Schutzbereich und Schranke auf Grundrechtsebene und im UrhG vgl. die Ausführungen auf S. 62 ff.

263 *Spindler* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 51 Rn. 65; ebenso *Rieger,* Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 42.

264 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60c Rn. 5; *Stieper* in Loewenheim UrhR-HdB, § 35 Rn. 27.

265 *A. Nordemann* in Fromm/Nordemann, UrhG, § 60c Rn. 4.

266 *A. Nordemann* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60c Rn. 4; *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 60c Rn. 1; *A. Nordemann* in Fromm/Nordemann, UrhG, § 60c Rn. 4.

267 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60c Rn. 5; *de la Durantaye,* Allgemeine Bildungs- und Wissenschaftsschranke, S. 55.

268 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60c Rn. 5; *Stieper* in Loewenheim UrhR-HdB, § 35 Rn. 27; *Ruffert* in Calliess/Ruffert, EUV/AEUV,

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

kenbestimmungen und den unionsrechtlichen Richtlinien dahingehend, ob die wissenschaftliche Tätigkeit kommerziell oder nicht-kommerziell erfolgt, ist auf übergeordnete und sich in der unionsrechtlichen Ermächtigungsnorm niederschlagenden[269] Wertungen zurückzuführen, die den Rechteinhabern fremde Verwertungen ihres Geisteswerks oder ihrer anderen Schutzgegenstände außerhalb vertraglicher Vereinbarungen nur dann zumuten, wenn sie keine kommerziellen Zwecke verfolgen.

III. Einordnung des Text- und Data-Minings

Der Frage, ob TDM stets als wissenschaftliche bzw. forschende Tätigkeit einzuordnen ist, sei es im primärrechtlichen Sinne oder im Sinne des deutschen GGs oder der GRCh, ist differenziert zu begegnen.

Das TDM kann, wie bereits deutlich wurde, verschiedenen Zwecken dienen. Es kann als Forschungsmethode, etwa in den digitalen Geisteswissenschaften, genutzt werden, aber auch zu journalistischen Zwecken und in Unternehmen. Immer dann, wenn ein Erkenntnisgewinn angestrebt wird, ist nicht nur der urheberrechtliche Wissenschaftsbegriff, sondern auch das Wissenschaftsverständnis von GG und GRCh betroffen, das gilt auch im kommerziellen Bereich. Dasselbe kann für das Training im maschinellen Lernen zutreffen, wenn an KI-Systemen als solchen geforscht wird oder die trainierten Systeme anderweitig zu Wissenschaftszwecken genutzt werden sollen, denn dann handelt es sich um Vorbereitungshandlungen. Etwas anderes gilt für anwendungsbezogenes Training von KI-Systemen zur Pro-

---

6. Aufl. 2022, Art. 13 Rn. 6; *Jarass* in Jarass, Charta der Grundrechte der Europäischen Union Art. 13 Rn. 8; *I. Augsberg* in von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht, Art. 13 Rn. 5; *Pünder* in Ehlers, Europäische Grundrechte und Grundfreiheiten, § 18 Rn. 8; *C. Thiele* in Frankfurter Kommentar EUV, GRC, AEUV, Art. 13 GRCh Rn. 12 mit Rückgriff auf OECD, Frascati Manual 2002, S. 30 f., 93; Europäische Kommission, Empfehlung der Kommission vom 11. März 2005 über die Europäische Charta für Forscher und einen Verhaltenskodex für die Einstellung von Forschern (2005/251/EG) Abschnitt 3, der auf eben diese Definition verweist.

269 Gem. Art. 5 Abs. 3 Nr. 1 Alt. 2 InfoSoc-RL können die EU-Mitgliedsstaaten *„für Zwecke der wissenschaftlichen Forschung"* Ausnahmen oder Beschränkungen vom Recht der Vervielfältigung und der öffentlichen Zugänglichmachung vorsehen, vorsehen, *„sofern — außer in Fällen, in denen sich dies als unmöglich erweist — die Quelle, einschließlich des Namens des Urhebers, wann immer dies möglich ist, angegeben wird und soweit dies zur Verfolgung nicht kommerzieller Zwecke gerechtfertigt ist"*, s. auch *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 60c Rn. 4.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

duktentwicklung, bei dem es nicht auf die Erlangung neuer Erkenntnisse ankommt. Insofern stellt das TDM in den meisten Fällen eine Forschungstätigkeit dar, der Einzelfall bedarf aber einer genauen Analyse.[270]

B. Ein Recht des Urhebers?

Das deutsche und europäische Urheberrecht verfolgte seit jeher das Ziel, urheberrechtliches Werkschaffen durch ein möglichst hohes Schutzniveau zu schützen, d. h. mittels einer weitreichenden Kontrolle des Urhebers über das eigene Geisteswerk und dessen Verwertung. Inzwischen sind in diesem Kontext verschiedene Aufweichungen zu bemerken. Die Erläuterung dieser systematischen Fragestellungen und urheberrechtlicher Grundlagen sind relevant für die spätere Einordnung, durch welche urheberrechtlichen Mechanismen wissenschaftliche Aktivität betroffen sein kann, einerseits auf Ebene des Schutzes ihrer Erzeugnisse, andererseits in der Ermöglichung der Nutzbarkeit anderer Werke oder Schutzgegenstände.

I. Der Urheber als Zentrum des Urheberrechts und Tendenzen zur Aufweichung

Das Ziel des traditionellen Urheberrechts, den Urheber in größtmöglichem Umfang zu schützen, zeigte sich einerseits in einer Absenkung der Schutzvoraussetzungen und andererseits in einer restriktiven Ermöglichung der Nutzung urheberrechtlicher Werke durch Dritte. Einen bedeutenden Einfluss auf die Absenkung der urheberrechtlichen Schutzvoraussetzungen haben dabei die weitreichende unionsrechtliche Harmonisierung des Urheberrechts und die Einflüsse der Rechtsprechung des EuGH: Zwar findet der Werkbegriff als solcher nur in einzelnen Werkkategorien Anklang – den Computerprogrammen, den Lichtbildwerken sowie den Datenbankwerken – und ist nicht ausdrücklich europarechtlich harmoni-

270 Anderer Auffassung ist *Stieper*, demzufolge das TDM generell Forschung „*im weitesten Sinne*" darstellt, weil jede automatisierte Analyse von Werken, die Informationen gewinnen soll – unabhängig von der Zweckbestimmung – Forschung sei, *Stieper*, ZUM 2021, S. 776, 777.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

siert.[271] Das in Bezug auf die genannten Schutzrechte gebildete Erfordernis einer „*eigenen geistigen Schöpfung*" wird durch den EuGH allerdings einheitlich ausgelegt[272] und wurde durch „*freie kreative Entscheidungen*"[273] fortgebildet.[274]

Über die Anerkennung eines Teilwerkschutzes in Art. 2 InfoSoc-RL[275] und unionsautonomer Auslegung des Vervielfältigungsrechts hat der EuGH sein Kriterium eines Werks auf andere Werkarten ausgeweitet[276] und dadurch ein unionsrechtliches Verständnis urheberrechtlicher Werke geprägt, weswegen nach umstrittener Ansicht ein einheitliches Verständnis des Werks gilt.[277]

---

271 Art. 1 Abs. 3 Computerprogramm-RL, Art. 6 Schutzdauer-RL, Art. 3 Abs. 1 Datenbank-RL.

272 *Handig*, GRUR Int. 2012, S. 973, 974; Rn. 35; EuGH, ECLI:EU:C:2009:465 - *Infopaq* = GRUR 2009, S. 1041, 1044; EuGH, ECLI:EU:C:2011:631 - *Football Association Premier League u. Murphy* = GRUR 2012, S. 156, 165 Rn. 188; EuGH, ECLI:EU:C:2012:115 - *Football Dacato/Yahoo* = GRUR 2012, S. 386, 387 Rn. 29 ff.

273 *Handig*, GRUR Int. 2012, S. 973 ff. in Anspielung auf das durch den EuGH gebildete Kriterium, das der Originalität Konturen verleihen soll, etwa EuGH, ECLI:EU:C:2011:798 - *Painer/Standard* = GRUR 2012, S. 166, 186 Rn. 89.

274 *Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 4 mit Verweis auf EuGH, ECLI:EU:C:2010:816 - *BSA/Kulturministerium* = GRUR 2011, S. 220, 222 Rn. 45 ff.; EuGH, ECLI:EU:C:2011:798 - *Painer/Standard* = GRUR 2012, S. 166, 168 Rn. 87 ff.; EuGH, ECLI:EU:C:2011:631 - *Football Association Premier League u. Murphy* = GRUR 2012, 156, S. 160 Rn. 96 ff.; EuGH, ECLI:EU:C:2012:115 - *Football Dacato/Yahoo* = GRUR 2012, S. 386, 388 Rn. 37 ff.; EuGH, ECLI:EU:C:2014:25 - *Nintendo Unternehmen/PC Box ua* = GRUR 2014, S. 255, 256 Rn. 45 ff.; ebenso *Schack*, GRUR 2021, S. 904, 904; *Schack,* Urheber- und Urhebervertragsrecht Rn. 187, demzufolge der EuGH „*ungeniert*" einen europäischen Werkbegriff für sich beanspruche.

275 Zur Absenkung der urheberrechtlichen Schutzvoraussetzungen sei insbesondere die *Infopaq*-Entscheidung aus dem Jahr 2009 erwähnt, in der der Gerichtshof auch schon Werkteilen, bestehend aus elf Wörtern, eine urheberrechtliche Schutzfähigkeit zubilligte, EuGH, ECLI:EU:C:2009:465 – *Infopaq* = GRUR 2009, S. 1041.

276 EuGH, ECLI:EU:C:2009:465 – *Infopaq* = GRUR 2009, S. 1041, 1044 Rn. 37; EuGH, ECLI:EU:C:2010:816 – *BSA/Kulturministerium* = GRUR 2011, S. 220, 222 Rn. 45.

277 Dafür u. a. *Handig*, GRUR Int. 2012, S. 973, 979; *Leistner*, GRUR 2019, S. 1114, 1114 f.; *Leistner*, GRUR 2014, S. 1145, 1145 f.; *Leistner*, ZGE 2013, S. 4, 43; *Leistner*, JZ 2011, S. 1140, 1144; eingeschränkt auch *Ohly,* Urheberrecht in der digitalen Welt - Brauchen wir neue Regelungen zum Urheberrecht und dessen Durchsetzung?, S. 31 f.; *Kriesel,* Einheitlicher europäischer Werkbegriff und Herabsenkung der Anforderungen an die Gestaltungshöhe bei Werken der angewandten Kunst, S. 161 f.; *J. H. Schmidt,* Maximalschutz im internationa-

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Die restriktive Zulassung der Nutzung urheberrechtlicher Werke wiederum wird insbesondere anhand der InfoSoc-RL sichtbar, mit der ausdrücklich bezweckt wurde, ein hohes Schutzniveau in den Mitgliedsstaaten zu gewährleisten,[278] indem sie Ausnahmen oder Beschränkungen abschließend aufzählt und sie weitestgehend freiwillig statt verpflichtend vorsieht. Nur die Erlaubnis der temporären Vervielfältigung (Art. 5 Abs. 1 InfoSoc-RL), die notwendigerweise in jedem digitalen Prozess erfolgen, ist verpflichtend vorgesehen. Diese verpflichtende Zulassung ist insofern das Mindestmaß eines Rechtsrahmens, der die Funktionsfähigkeit des Internets ermöglicht.

Der Ansatz, Schrankenbestimmungen aufgrund eines – wie nachfolgend erläutert wird nur vermeintlichen – Regel-Ausnahme-Verhältnisses restriktiv auszulegen,[279] trug wiederum auf Ebene der Rechtsauslegung dazu bei, dass die Verwendung fremder Werke auch innerhalb gesetzlicher Schran-

---

len und europäischen Urheberrecht, S. 151 f.; *Roder,* Die Methodik des EuGH im Urheberrecht, S. 370; *Loewenheim/Leistner* in Loewenheim UrhR-HdB, § 6 Rn. 5 f.; *A. Nordemann* in Fromm/Nordemann, UrhG, § 2 Rn. 7a mit Ausahme der angewandten Kunst; ebenso *A. Nordemann* in FS Bornkamm, S. 895, 898 f.; tendenziell auch *Bisges*, ZUM 2015, S. 357, 359 f.; in begrenztem Umfang unter Beibehaltung eines „*Interpretationsspielraums*" der mitgliedsstaatlichen Gerichte *Steinbeck*, EuZW 2014, S. 329, 329 f.; nicht ausdrücklich Position ergreifen dagegen *Bullinger* in Wandtke/Bullinger, UrhG, 5. Aufl. 2019, § 2 Rn. 14; ablehnend *E.-M. König,* Der Werkbegriff in Europa, S. 26 ff.

278 InfoSoc-RL Erw. 4.

279 Bereits das RG sprach von „*Ausnahmen, die die Regel des ausschließlichen Urheberrechts durchbrechen*", RG, Urt. v. 26.03.1930, I 260/29 - *Schlager-Liederbuch* = RGZ 128, 102, S. 104. Lange Zeit wurde dieser Grundsatz vom BGH rezipiert: BGH, Urt. v. 03.04.1968 - I ZR 83/66 – *Kandinsky* = GRUR 1968, S. 607, 608: „*Soweit der Urheber […] Beschränkungen seiner ausschließlichen Verwertungsrechte dulden muß, handelt es sich aber um eine Ausnahme […]. Die Ausnahmevorschrift des § 51 Nr. 1 UrhG ist in diesen durch ihren Sinn und Zweck gebotenen Grenzen auszulegen*". Ähnlich BGH, Urt. v. 23.01.1964 – III ZR 12/63 – *Landesversicherungsanstalt* = GRUR 1972, S. 614, 615: „*Eine enge Auslegung […] ist schon deshalb geboten, weil Art. 11 RBÜ (Brüsseler Fassung) dem Komponisten bezüglich öffentlicher Aufführungen ein uneingeschränktes ausschließliches Recht gewährt und dem deutschen Gesetzgeber daher nur ein kleiner Spielraum für die Zulassung eng begrenzter Ausnahmen verblieb*"; genauso BGH, Urt. v. 17.03.1983 - I ZR 186/80 – *Zoll- und Finanzschulen* = GRUR 1983, S. 562, 653; BGH, Urt. v. 18.04.1985 - I ZR 24/83 - *Schulfunksendung* = GRUR 1985, S. 874, 275; BGH, Urt. v. 06.06.1991 - I ZR 26/90- *Liedersammlung* = GRUR 1991, S. 903, 705; auch noch im Jahr 2012 BGH, Urt. v. 30. 11. 2011 – I ZR 212/10 - *Blühende Landschaften* = GRUR 2012, S. 819, 822 Rn. 28: „*In diesem Zusammenhang ist zu berücksichtigen, dass die auf der Sozialbindung des geistigen Eigentums beruhenden Schrankenbestimmungen […] generell eng auszulegen sind*".

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

kenbestimmungen nur in eng begrenztem Umfang möglich war. Dazu tendierte auch der EuGH.[280]

Angesichts einiger Reformen lässt sich die Tendenz zu einer Abkehr vom urheberzentrierten Dogma erkennen, das vor allem die Zulassung und Auslegung der Schranken betrifft. In der rechtswissenschaftlichen Literatur wird der sog. Grundsatz der engen Schrankenauslegung kritisch gesehen und stattdessen eine Auslegung entsprechend des Sinns und Zwecks der Schranken vorgenommen.[281] Begründet wird diese nun weitestgehend vorherrschende Auffassung damit, dass es im deutschen Recht keinen Grundsatz gibt, nach dem Ausnahmen stets restriktiv auszulegen[282] und

---

280 EuGH, ECLI:EU:C:2009:465 – *Infopaq* = GRUR 2009, S. 1041, 1045 Rn. 56: „*Für die Auslegung der einzelnen Voraussetzungen ist sodann daran zu erinnern, dass nach ständiger Rechtsprechung die Bestimmungen einer Richtlinie, die von einem in dieser Richtlinie aufgestellten allgemeinen Grundsatz abweichen, eng auszulegen sind*". Mit der gleichen Begründung u. a. EuGH, ECLI:EU:C:2012:16 – *Infopaq II* = GRUR Int 2012, 336, S. 339 Rn. 27; EuGH, ECLI:EU:C:2014:1195 – *Public Relations Consultants Association* = GRUR 2014, . 654, 655 Rn. 23; EuGH, ECLI:EU:C:2012:65 – *Luksan* = GRUR 2012, S. 489, 494 Rn. 101. Beinahe noch deutlicher äußerte sich der EuGH in EuGH, ECLI:EU:C:2014:254 – *ACI Adam u.a.* = GRUR 2014, S. 546, 547 Rn. 22 f., 27: „*Überdies geht aus dem 44. Erwägungsgrund […] hervor, dass es die Absicht des Unionsgesetzgebers war, den Umfang der […] Ausnahmen oder Beschränkungen […] möglicherweise noch enger zu begrenzen. Hingegen sehen weder dieser Erwägungsgrund noch irgendeine andere Bestimmung dieser Richtlinie die Möglichkeit vor, dass die Mitgliedstaaten den Umfang solcher Beschränkungen und Ausnahmen ausweiten*".

281 U. a. *T. Dreier*, JZ 2003, S. 477, 478; *Geiger* in: Hilty/Peukert (Hrsg.), Interessenausgleich im Urheberrecht, S. 143, 152; *P. Raue* in FS Nordemann, S. 327, 339; *Rehse,* Zulässigkeit und Grenzen ungeschriebener Schranken des Urheberrechts, S. 61; *Poeppel,* Die Neuordnung der urheberrechtlichen Schranken im digitalen Umfeld, S. 45 f.; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 65 ff.; *Stieper*, ZUM 2019, S. 713, 715; *Wandtke/R. König*, ZUM 2014, S. 921, 928 f.; *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 36 f.; letztlich auch *Findeisen,* Die Auslegung urheberrechtlicher Schrankenbestimmungen, S. 73 ff.; explizit hinsichtlich digitalen Nutzungshandlungen *Kröger*, MMR 2002, S. 18, 19 f.; von einem strikten Regel-Ausnahmeverhältnis geht *Förster,* Fair Use, S. 183 aus; skeptisch ebenso *Schunke* in FS Wandtke, S. 341, 345; *Schack* in FS Schricker, S. 511, 514 f. mit dem Zusatz, dass dennoch analoge Anwendungen möglich seien auf den alten Standpunkt stellen sich *Dreyer* in Dreyer/Kotthoff, UrhG, § 2 Rn. 25 mit Veweis auf die o. g. Rechtsprechung *Schack,* Urheber- und Urhebervertragsrecht Rn. 98, 563.

282 Zu den Schranken *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Vor § 44a Rn. 7; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 66; *Kleinemenke,* Fair Use im deutschen und europäischen

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Schranken überdies bereits aus dogmatischer Sicht schon nicht grundsätzlich als Ausnahmen, sondern vielmehr als inhaltliche Beschränkungen zu klassifizieren sind.[283] Außerdem trägt gerade die Sozialpflichtigkeit des Eigentums bzw. die Zulassung von urheberrechtlichen Schranken den Interessen der Allgemeinheit Rechnung.[284]

In der Rechtsprechung ist jedenfalls eine leichte Erweiterung der Schrankenauslegung zu bemerken, so entschied BVerfG bereits im Jahr 1986, dass eine enge Schrankenauslegung u. U. nicht mit den Grundrechten zu vereinbaren ist.[285] Auch der BGH hatte bereits zwischenzeitlich eine Auslegung der Schranken nach ihrem Sinn und Zweck vorgenommen.[286] Zudem hat auch der EuGH immer wieder Tendenzen geäußert, die dem vormals vertretenen Grundsatz der engen Schrankenauslegung entgegen-

---

Urheberrecht?, S. 78; *Hilty* in FS Schricker, S. 325, 327 f.; *Poeppel,* Die Neuordnung der urheberrechtlichen Schranken im digitalen Umfeld, S. 43; *Schack* in FS Schricker, S. 511, 514. Aus rechtstechnischer Perspektive ist das allgemein anerkannt, vgl. nur *Bydlinski,* Juristische Methodenlehre und Rechtsbegriff, S. 440 f.; in Bezug auf die analoge Anwendung von Ausnahmevorschriften grundlegend *Canaris,* Die Feststellung von Lücken im Gesetz, S. 181f.; für die Zulässigkeit einer analogen Anwendung von Ausnahmevorschriften auch *Engisch,* Einführung in das juristische Denken, S. 211; *Larenz/Canaris,* Methodenlehre der Rechtswissenschaft, S. 174 ff.; *F. Müller/Christensen,* Grundlegung für die Arbeitsmethoden der Rechtspraxis, S. 399 Rn. 373.

283 Vgl. dazu die nachfolgenden Erläuterungen, S. 61 ff.

284 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 36; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 67 f.

285 *Schunke* in FS Wandtke, S. 341, 344; BVerfG, Beschl. v. 17. 11. 2011 - 1 BvR 1145/11 – *Kunstausstellung im Online-Archiv* = GRUR 2012, 389, S. 319 Rn. 17.

286 Demzufolge soll der Grundsatz der engen Schrankenauslegung einer Auslegung nach Sinn und Zweck solange nicht entgegenstehen, wie eine Gesetzeslücke existiert und eine Analogie geboten erscheint, BGH, Urt. v. 04.12.1986 - I ZR 189/84 – *Filmzitat* = GRUR 1987, S. 362, 363.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

stehen oder ihn jedenfalls aufweichen,[287] obwohl er zwischenzeitlich auch restriktivere Urteile fällte.[288]

Daneben sind die Schrankenbestimmungen auch durch gesetzgeberische Wertungen gestärkt worden. Die DSM-RL aus dem Jahr 2019 sollte zwar das durch die anderen Richtlinien gesicherte hohe Schutzniveau beibehalten, insbesondere im Bereich der Erhaltung des Kulturerbes, der Bildung, Wissenschaft und Innovation[289] sollten jedoch *„Anreize für Innovation, Kreativität, Investitionen und die Produktion neuer Inhalte, auch im digitalen Umfeld"*[290] gesetzt werden. Das zeigt sich an der Ausgestaltung der Schrankenbestimmungen: Bestimmte Ausnahmen oder Beschränkungen wurden verpflichtend, vgl. Art. 3 – Art. 6 DSM-RL, vertragsfest, Art. 7 Abs. 1 DSM-RL, sowie im digitalen Raum gegenüber technischen Schutzmaßnahmen beständig ausgestaltet, Art. 7 Abs. 2 DSM-RL.[291]

Eine wichtige Einschränkung der Schrankenauslegung stellt weiterhin der auf völkerrechtlichen Vorgaben basierende[292] Drei-Stufen-Test dar, der in Art. 5 Abs. 5 InfoSoc-RL ausdrücklich kodifiziert wurde und als Schranken-Schranke[293] nicht nur für die Schaffung von Schranken, sondern auch für ihre richterliche Rechtsauslegung zu beachten ist.[294]

---

287 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 939 Rn. 69 ff.; EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, S. 940, 944 Rn. 53 ff.; auf die praktische Wirksamkeit verweist auch EuGH, ECLI:EU:C:2014:2196 – *Eugen Ulmer* = GRUR 2014, S. 1078, 1079 Rn. 32; EuGH, ECLI:EU:C:2011:798 – *Painer/Standard* = GRUR 2012, S. 166, 171 Rn. 133 beharrt noch auf der restriktiven Auslegung, möchte aber eine praktische Wirksamkeit gewährleisten; ebenso EuGH, ECLI:EU:C:2011:631 – *Football Association Premier League u. Murphy* = GRUR 2012, S. 156, 164 Rn. 166 f.; EuGH, ECLI:EU:C:2014:2132 – *Deckmyn* = GRUR 2014, S. 972, 973 Rn. 22 f.

288 *Wandtke/R. König*, ZUM 2014, S. 921, 928 f.; EuGH, ECLI:EU:C:2014:254 – *ACI Adam u.a.* = GRUR 2014, S. 546, 547 Rn. 22 f. mit Verweis auf; EuGH, ECLI:EU:C:2009:465 – *Infopaq* = GRUR 2009, S. 1041, 1045 Rn. 56.

289 DSM-RL Erw. 3, 5.

290 DSM-RL Erw. 2.

291 In diesen digitalen Systemen setzten sich die gesetzlichen Schranken nur eingeschränkt gegen die Rechteverwaltung durch, § 95b Abs. 3 UrhG a. F. bzw. Art. 6 Abs. 4 UAbs. 4 InfoSoc-RL, dazu ausführlich *Specht-Riemenschneider,* Diktat der Technik, S. 395 ff. Zur neuen Rechtslage im nationalen Umsetzungsrecht vgl. die Erläuterungen an späterer Stelle, S. f.

292 Art. 10 Abs. 1 WCT und Art. 16 Abs. 2 WPPT; dazu *v. Welser* in Wandtke/Bullinger, UrhG, 5. Aufl. 2019, § 44a Rn. 22.

293 *Senftleben*, GRUR Int. 2004, S. 200, 200.

294 *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 73; *Geiger*, IIC 2006, S. 74, 80; BGH, Urt. v. 11. 7. 2002 - I ZR

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

II. Die systematische Stellung von Schrankenbestimmungen

Dass Schrankenbestimmungen systematisch nicht als Ausnahme von dem Schutzrecht betrachtet werden können, zeigt auch die systematische Stellung der Schrankenbestimmungen auf Grundrechtsebene im GG und der GRCh sowie auf primärrechtlicher Ebene.

1. Das Verhältnis von geistigem Eigentum und Sozialbindung imdeutschen und europäischen Grundrechtsgefüge

Als privates Vermögensrecht fallen die vermögenswerten Elemente des geistigen Eigentums und damit auch der Schutzrechte des UrhGs unter den Schutzbereich der Eigentumsgarantie des deutschen Grundgesetzes, Art. 14 Abs. 1 S. 1 GG.[295] Die einfachrechtlichen Schranken bilden demgegenüber als abstrakt-generelle Vorgaben hinsichtlich des Umgangs mit dem Eigentum, die nicht auf die Substanz des Eigentums zugreifen,[296] auf der Sozialpflichtigkeit des Eigentums basierende Inhalts- und Schrankenbestimmungen[297] gem. Art. 14 Abs. 1 S. 2 GG.[298] Aus systematischer Perspektive bestimmt der Gesetzgeber mit den Inhalts- und Schrankenbestimmungen erst den Inhalt des Eigentums,[299] wobei er eine Verhältnis-

---

255/00 – *Elektronischer Pressespiegel* = GRUR 2002, S. 963, 967; BGH, Urt. v. 16.01.1997 - I ZR 9/95 – *CB-Infobank I* = GRUR 1997, S. 459, 463 mit Verweis auf die RBÜ in der Pariser Fassung.

295 *Papier/Shirvani* in Dürig/Herzog/Scholz, GG, Art. 14 Rn. 314 f.; BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66 – *Kirchen- und Schulgebrauch* = GRUR 1972, S. 481, 483; BVerfG, Beschl. v. 11.10.1988 - 1 BvR 743/86 = NJW 1992, 1307, S. 1308; BVerfG, Beschl. v. 03.10.1989 - 1 BvR 775/86 – *Vermietungsvorbehalt* = GRUR 1990, S. 183, 184.

296 *Papier/Shirvani* in Dürig/Herzog/Scholz, GG, Art. 14 Rn. 417.

297 Deren Differenzierung erübrigt sich, *Papier/Shirvani* in Dürig/Herzog/Scholz, GG, Art. 14 Rn. 417; BVerfG, Urt. v. 1. 3. 1979 - 1 BvR 532, 533/77, 419/78, 1 BvL 21/78 – *Verfassungsmäßigkeit des Mitbestimmungsgesetzes* = NJW 1979, S. 699, 703; BVerfG, Besch. v. 14.07.1981 - 1 BvL 24/78 – *Verfassungsrechtliche Grenzen für Pflichtexemplarregelungen* = NJW 1982, S. 633, 633 f.

298 *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 42 ff.; *Fechner,* Geistiges Eigentum und Verfassung, S. 199.

299 BVerfG, Urt. v. 1. 3. 1979 - 1 BvR 532, 533/77, 419/78, 1 BvL 21/78 – *Verfassungsmäßigkeit des Mitbestimmungsgesetzes* = NJW 1979, S. 699, 703; BVerfG, Beschl. v. 07.07.1971 - 1 BvR 765/66 – *Kirchen- und Schulgebrauch* = GRUR 1972, S. 481, 483; BVerfG, Urt. v. 18.12.1968 - 1 BvR 638, 673/64, 200, 238, 249/56 – *Grenzen der Legalenteignung* = NJW 1969, S. 309, 309 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

mäßigkeit wahren und in Fällen unzumutbarer Härte einen finanziellen Ausgleich berücksichtigen muss.[300]

Auch im europäischen Verständnis wird der Inhalt des Eigentums, Art. 17 Abs. 2, 1 GRCh, dem das Urheberrecht als Teilbereich des geistigen Eigentums unterfällt, erst durch den Gesetzgeber bestimmt.[301] Regelungen, die nicht auf die Substanz des Eigentums zugreifen, sondern den Umgang mit ihm bestimmen, sind ebenso wie aus deutscher Perspektive auch aus europäischer Perspektive Nutzungsregelungen gem. Art. 17 Abs. 1 S. 3 GRCh, die einer Überprüfung unterliegen, ob sie für das Wohl der Allgemeinheit erforderlich sind.[302] Insofern ist eine Parallele zur Sozialpflichtigkeit nach deutschem Verfassungsrecht festzustellen.[303] Durch die Vorgabe, dass die Nutzungsregelungen den Inhalt des Eigentums bestimmen, ist die Eigentumsordnung auch nach EU-Recht mitgliedsstaatlich geregelt (vgl. auch Art. 345 AEUV),[304] wobei das EU-Recht durch Verordnungen und Richtlinien gewisse Rahmenbedingungen setzt und Einflüsse auf die nationalstaatliche Eigentumsordnung nimmt.

2. Das Verhältnis von Werkschutz und Schranke im einfachen Recht

Auch die Frage, wie sich Schrankenbestimmungen im deutschen Primärrecht zum Urheberrecht verhalten, wird mittlerweile weitestgehend so beurteilt, dass die durch die Schrankenbestimmungen erlaubten Nutzungs-

---

300 *Fechner,* Geistiges Eigentum und Verfassung, S. 230 ff.; BVerfG, Besch. v. 14.07.1981 - 1 BvL 24/78 – *Verfassungsrechtliche Grenzen für Pflichtexemplarregelungen* = NJW 1982, S. 633. Die gem. Art. 5 Abs. 3 S. 1 GG ihrerseits verfassungsrechtlich geschützte wissenschaftliche Forschung ist ein im Rahmen der Verhältnismäßigkeit zu berücksichtigender legitimer Zweck solcher Gesetze, die den Inhalt des Eigentumsgrundrechts als Inhalts- und Schrankenbestimmungen definieren.

301 *N. Bernsdorff* in Meyer/Hölscheidt, EU GRCh, Art. 17 GRCh Rn. 13; *Calliess* in Calliess/Ruffert, EUV/AEUV, Art. 17 GRCh Rn. 4; ebenso *Vosgerau* in Stern/Sachs, GRCh, Art. 17 Rn. 58; *Callies* in Ehlers, Europäische Grundrechte und Grundfreiheiten, § 20 Rn. 15: „*normgeprägter Schutzbereich*"; kritisch sieht das *F. Wollenschläger* in von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht, Art. 17 GRCh Rn. 22, der auch eigentumsdefinierende Normen als Beschränkungen des Eigentums für die Zukunft versteht.

302 *N. Bernsdorff* in Meyer/Hölscheidt, EU GRCh, Art. 17 GRCh Rn. 20; *Jarass*, NVwZ 2006, S. 1089, 1093.

303 *Calliess* in Calliess/Ruffert, EUV/AEUV, Art. 17 GRCh Rn. 13.

304 *Calliess* in Calliess/Ruffert, EUV/AEUV, Art. 17 GRCh Rn. 4; *Callies* in FS Fiedler, S. 463, 475.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

handlungen systematisch vom Schutzbereich des UrhG ausgenommen werden, sie dem geistigen Eigentum also nie zugeordnet waren.[305] Die Schranken bestimmen also erst die Reichweite des Urheberrechts.[306] Das basiert einerseits auf der sich immer stärker durchsetzenden Auffassung, dass aus der gesetzlichen Systematik kein generelles Regel-Ausnahme-Verhältnis von Werkschutz und Nutzungserlaubnis zu schlussfolgern ist[307] und wird andererseits auf die zuvor dargestellten Grundrechtswertungen gestützt:[308] Der verfassungsrechtliche Grundsatz, dass das primäre Recht

---

305 U. a. *T. Dreier*, GRUR Int. 2015, S. 648, 651; *Geiger* in: Hilty/Peukert (Hrsg.), Interessenausgleich im Urheberrecht, S. 143, 151; *Haß* in FS Klaka, S. 127, 133 bezeichnet die Vorstellung, der Gesetzgeber räume ein allumfassendes Recht ein, das erst anschließend beschränkt würde, als ; *Rehse,* Zulässigkeit und Grenzen ungeschriebener Schranken des Urheberrechts, S. 49; *Schack*, GRUR 2021, S. 904 ff.; *Schack,* Urheber- und Urhebervertragsrecht Rn. 540 behandelt die zeitlichen und inhaltlichen Schranken insofern gleich; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 130 f.; *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 34 *„gekünstelt"*; *Findeisen* hingegen differenziert zwischen den Zielsetzungen, die den Schranken zugrunde liegen. Nur solche aus dem Wesen des Urheberrechts und der geistigen Auseinandersetzung resultierende Schranken gelten demnach als Beschränkungen des Rechts, *Findeisen,* Die Auslegung urheberrechtlicher Schrankenbestimmungen, S. 100 ff., 113; gegen eine Herleitung von Eigenschaften aus dem Begriff der Schranke äußern sich *Zurth,* Rechtsgeschäftliche und gesetzliche Nutzungsrechte im Urheberrecht, S. 98 ff.; *Schack* in FS Schricker, S. 511, 511; nicht eindeutig *Seith,* Wie kommt der Urheber zu seinem Recht?, S. 10 ff.

306 *Berger*, GRUR 2017, S. 953, 954; jedenfalls in Bezug auf das subjektive Urheberrecht *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 1; *Geiger*, GRUR Int. 2008, S. 459, 461. In diesem Sinne vertritt *Stieper* sogar, dass urheberrechtliche Schranken eine Position des Nutzers manifestieren, die sich zwar nicht bzw. nicht zwingend in einer Anspruchsposition äußert, doch über die Vertragsfestigkeit eine subjektive Rechtsposition verleihen, *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 165 f., 429 f.

307 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 34; vgl. dazu die Ausführungen auf S. 79 ff.

308 In diesem Kontext ist dennoch einzuräumen, dass auch die Sozialbindung des Urheberrechts philosophisch jedenfalls in einem Ansatz damit begründet wird, dass sich neues Werkschaffen auf bestehendem aufbaut, vgl. die ausführlichen Darstellungen von *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 23 ff. Von heutiger verfassungsrechtlicher Ebene lässt sich dies wohl nicht halten. Passender erscheint in diesem Zusammenhang der (ebenfalls von *Stieper* dargestellte) Begründungsansatz, dass Werkschaffen bestimmungsgemäß auch der Gesellschaft (in diesem Sinne sowohl Allgemeininteressen als auch bestimmte Individualinteressen) nutzt, der auch

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

das Eigentum erst bestimmt, ist bereits in § 1 UrhG enthalten, dort heißt es, dass Urheber den Schutz *„nach Maßgabe dieses Gesetzes"* erhalten.[309] Nach deutscher Systematik handelt es sich bei den Schrankenbestimmungen demnach um Beschränkungen eines Rechts statt um Ausnahmen.

Im Unionsrecht ist das uneindeutig, so wird zwischen Ausnahmen und Beschränkungen unterschieden, deren systematische Einordnung sich von der bisher dargelegten Einordnung unterscheidet. Die Richtlinien implizieren dabei ein Rangverhältnis, das Ausnahmen den Beschränkungen überordnet: Die InfoSoc-RL legt für die fakultativen Nutzungserlaubnisse nach Art. 5 Abs. 2 und 3 InfoSoc-RL nur fest, dass es sich um Ausnahmen oder Beschränkungen handeln solle und lässt den Mitgliedsstaaten offen, wie stark oder schwach sie diese freiwilligen Nutzungserlaubnisse gesetzessystematisch ausgestalten. Der verpflichtende Art. 5 Abs. 1 InfoSoc-RL, der im deutschen Recht in § 44a UrhG umgesetzt wurde, ist jedoch als Ausnahme ausgestaltet. Das deutet auf eine wertungsmäßige Privilegierung hin. Dasselbe gilt in Bezug auf Art. 3 DSM-RL, die verpflichtende Erlaubnis wissenschaftlichen, nicht-kommerziellen TDM, auch diese soll explizit als Ausnahme ausgestaltet werden, was als Privilegierung der Digitalforschung, d. h. als systematische Stärkung, zu verstehen ist.[310] Im Gegensatz dazu spricht der fakultative Art. 4 Abs. 1 DSM-RL von Ausnahmen oder Beschränkungen. Aufgrund dieser begrifflichen Unterschiede ist anzunehmen, dass das Unionsrecht Ausnahmen stärker einordnet als Beschränkungen, was jedoch kein grundsätzlich anderes Verständnis von Urheberrecht und Schrankenbestimmung begründet.[311]

---

vom deutschen Gesetzgeber aufgegriffen wurde, BT-Drs. IV/270, S. 62; BT-Drs. 10/837, S. 9.

309 *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 34; *Stieper*, GRUR 2020, S. 1, 131.

310 *B. Raue*, ZUM 2021, S. 793, 798; *Schack*, GRUR 2021, S. 904, 907 hingegen vermutet ein Redaktionsversehen.

311 Die Aussagekraft dieser Begriffe wird teilweise unterschiedlich erachtet. Nach bildlicher Darstellung von *Geiger* handelt es sich bei einer Ausnahme um eine Insel in einem Meer voller Freiheit, Beschränkungen sind hingegen eine Insel der Exklusivität in einem Meer voller Freiheiten, *Geiger* in: Hilty/Peukert (Hrsg.), Interessenausgleich im Urheberrecht, S. 143, 150; *Geiger*, GRUR Int. 2004, S. 815, 818 f. mit Verweis auf die französische Sichtweise; der gleichen Auffassung ist *T. Dreier*, ZUM 2013, S. 769, 773; *Dusollier*, IIC 2003, S. 62, 73; *Voorhof* in: Ginsburg/Besek (Hrsg.), Ad-juncts and Alternatives to Copyright, Proceedings of the ALAI Congress 2001, S. 636, 639; *Stieper* betrachtet die Unterscheidung hingegen als eine *„Frage der Gesetzestechnik"*, *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a Rn. 27.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

C. Die Besonderheiten des Wissenschaftsurheberrechts

Dem Urheberrecht kommt insgesamt eine erhebliche Steuerungswirkung bzgl. der Wissenschaft und Forschung zu, die bestimmten Regelungsmechanismen geschuldet ist. Teilweise hat sich für diesen Spezialbereich der Begriff des Wissenschaftsurheberrechts herausgebildet,[312] der auch in dieser Untersuchung gebraucht wird. Diesbezüglich liegt der Untersuchung das Verständnis zugrunde, dass der Begriff des Wissenschaftsurheberrechts gerade diejenigen Mechanismen meint, die sich in Bezug auf den wissenschaftlichen Urheber von der sonstigen Regelungsweise unterscheiden.[313] Das betrifft spezifische Ausprägungen urheberrechtlicher Grundwertungen sowie die besonderen Bedürfnisse, die, statt in einem hohen Schutzniveau gerade in erweiterten Schrankenbestimmungen, die eine Nutzbarkeit fremder Werke als Voraussetzung, wissenschaftlich zu arbeiten, anerkennt.

I. Das urheberrechtliche Schutzniveau in Wissenschaftskontexten

In Wissenschaftskontexten kommen einerseits klassischerweise Publikationen in Textform oder in grafischen Darstellungen zustande, die Forschungsansätze oder -ergebnisse enthalten. In jüngerer Zeit sind aber auch Forschungsdaten bzw. deren erhebliche Bedeutung stärker in den wissenschaftlichen Fokus gerückt.

---

312 Dieser Begriff ist kein feststehender Begriff, vgl. *Kuhlen*, Die Transformation der Informationsmärkte in Richtung Nutzungsfreiheit, S. 22, er wird aber immer häufiger gebraucht, z. B. von *Kuhlen* selbst, vgl. *Kuhlen*, IWP 2017, S. 227 ff.; sowie u. a. von *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Vor § 44a; *Peukert*, JIPITEC 2013, S. 142 ff., 5 Rn. 4; *Steinhauer*, GRUR-Prax 2014, S. 471 ff.; Max-Planck-Institut für Geistiges Eigentum, Wettbewerbs- und Steuerrecht, Stellungnahme des Max-Planck-Instituts für Geistiges Eigentum, Wettbewerbs- und Steuerrecht zur Anfrage des Bundesministeriums der Justiz vom 19. Februar 2009: Urheberrecht „Dritter Korb" (2009), S. 13, 18.

313 Vgl. auch *Wildgans*, Urheberrecht in der wissenschaftlichen Forschung: Ein "Window of Opportunity?", S. 57: „*Das Wissenschaftsurheberrecht bezeichnet im allgemeinen Sprachgebrauch die Gesamtheit der Rechtsnormen des deutschen Urheberrechtsgesetzes, die einen Bezug zur Wissenschaft im weiteren Sinne aufweisen*".

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

1. Ein eingegrenzter Schutz wissenschaftlicher Werke

Widersprüche zwischen kreativem Werkschaffen und analytischer Wissenschaft negativ auf das urheberrechtliche Schutzniveau aus, was auf verschiedene Gründe zurückzuführen ist. Das Kriterium des Werkschutzes, das für wissenschaftliche Texte teils schwierig zu erfüllen ist bzw. an dem ein Urheberrechtsschutz wissenschaftlicher Texte scheitern kann, ist das Kriterium der geistigen Schöpfung und der damit verbundenen notwendigen Individualität bzw. der sog. Schöpfungshöhe.[314] Wissenschaftliche Texte können schon ihrer Natur nach nur ein geringeres Maß an eigenschöpferischen Zügen aufweisen,[315] schließlich geht es um die Vermittlung wissenschaftlicher Erkenntnisse und gerade nicht darum, Empfindungen beim Leser zu wecken.[316] Zudem hängt das Schutzniveau wissenschaftlicher Werke maßgeblich davon ab, welche Elemente eines Werks überhaupt schutzfähig sind. Der historische Gesetzgeber bezog sowohl den Inhalt als auch die Form mit ein.[317] Dennoch wurde lange Zeit nach Form und Inhalt differenziert, d. h. nur die Form eines Werks wurde für schutzfähig eingeordnet, wobei weiter eine Unterscheidung der sog. inneren und äußeren Form vorgenommen wurde.[318]

---

314 *Krujatz,* Open Access, S. 68 f.

315 Mit Hinweis darauf, dass die Person bzw. der Urheber dann beliebig auswechselbar wäre *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 65 ff.; differenzierter betrachtet das *Ulmer,* Der Urheberschutz wissenschaftlicher Werke unter besonderer Berücksichtigung der Programme elektronischer Rechenanlagen, S. 16: „*Ein Urheberrechtsschutz ist insoweit nicht möglich, als die Sachgebundenheit die Freiheit individueller Gestaltung ausschließt. Der Schutz vermag aber da einzusetzen, wo die Vielheit möglicher Wege der Erläuterung oder der Anwendung der Erkenntnisse und Lehren beginnt und damit Raum für die Entfaltung des persönlichen Stils bleibt*".

316 *V. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 25; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022 § 2 Rn. 26: „*Enge Grenzen*".

317 BT-Drs. IV/270, S. 38; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 26.

318 So zuerst *Kohler,* Urheberrecht an Schriftwerken und Verlagsrecht, S. 143 ff. zum damaligen Autorenrecht; außerdem *Plander*, UFITA 1976/I, S. 25, 25: Die innere Form ist das System der Schilderung oder Erläuterung bzw. welcher Eindruck vom Geschilderten besteht, die äußere Form hingegen die konkrete Gestaltung, zur Freiheit wissenschaftlicher Inhalte statt fiktionaler bereits Ulmer, Urheber- und Verlagsrecht § 19 II; ebenso *Knap* in FS Troller, S. 117, 127; auf den Standpunkt der strikten Trennung stellen sich weiterhin *Dreyer* in Dreyer/Kotthoff, UrhG, § 2 Rn. 118.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Diese Trennung von Form und Inhalt stößt schon seit längerer Zeit auf Kritik,[319] Grund dafür sind im Wesentlichen Definitions-[320] und Abgrenzungsschwierigkeiten[321]: Schwierigkeiten warf insbesondere auf, dass die Rechtsprechung fiktionale Geschichten auch in inhaltlichen Elementen als schutzfähig erachtet, solange die Elemente ausreichend manifestiert sind.[322] Die rechtliche Beurteilung der Schutzfähigkeit lässt sich also nicht immer mit einem „*Nein*" hinsichtlich des Inhaltes und einem „*Ja*" hinsichtlich der Form beantworten. Außerdem zeigen sich auch andere Widersprüche, z. B. wurde die Gedankenführung wissenschaftlicher Wer-

---

319 Zuerst *Ulmer,* Urheber- und Verlagsrecht, S. 119 ff.; vgl. ausführlich *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 52, 95 ff.; *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 44 ff.; *Waiblinger,* "Plagiat" in der Wissenschaft, S. 56 ff.; *Götting* in FS Nordemann, S. 7, 9 ff.; zu den Besonderheiten nichtgegenständlicher Werke *Haberstumpf* in FS Schulze, S. 3 ff.; ebenso bereits *R. Köhler,* Der urheberrechtliche Schutz der Rechenprogramme, S. 53; *Hubmann* in FS Uchtenhagen, S. 175, 185 f. sieht aufgrund der Trennung von Form und Inhalt die Gefahr des Diebstahls wissenschaftlicher Erfolge; zu inhaltlichen Elementen, die sich in der Formgestaltung ausdrücken *W. Nordemann*, NJW 1970, S. 882, 882; *Schricker* in FS GRUR, S. 1095 ff., 1112 Rn. 43: *„Dass es im Interesse der Wissenschaftsfreiheit für den Urheberrechtsschutz wissenschaftlicher Erkenntnisse Grenzen geben muss, trifft sicher zu. Mit einer abstrakten begrifflichen Scheidung von Form und Inhalt ist das Problem aber nicht zu lösen*"; einen differenzierten Ansatz vertritt *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 44 ff.

320 *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 34.

321 *„Noch niemandem ist es gelungen, das Wesen der Idee von dem der Form adäquat abzugrenzen*", *Hoeren*, MMR 2000, S. 3, 3. Damit wendet er sich insbesondere auch gegen die Unterscheidung der sog. inneren und äußeren Form. Ebenfalls kritisch in Bezug auf die Trennung von Inhalt und Form sind *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 83; von *„fließenden Übergängen*" spricht *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 44.

322 *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 78; u. a. zu einer Fabel OLG Karlsruhe, Urt. v. 19.04.1956 - 4 U 96/55 – *Trotzkopf* = GRUR 1957, 395, S. 396 f.; BGH, Urt. v. 29.04.1999 - I ZR 65/96 – *Laras Tochter* = GRUR 1999, 984, S. 987 mit dem Verweis darauf, dass auch *„eigenpersönlich geprägte Bestandteile und formbildende Elemente des Werks, die im Gang der Handlung, in der Charakteristik und Rollenverteilung der handelnden Personen, der Ausgestaltung von Szenen und in der „Szenerie" eines Romans liegen*" urheberrechtlich geschützt sein können; zu einer Fabel OLG Karlsruhe, Urt. v. 19.04.1956 - 4 U 96/55 – *Trotzkopf* = GRUR 1957, 395, S. 396 f.; BGH, Urt. v. 1. 12. 2010 - I ZR 12/08 – *Perlentaucher* = GRUR 2011, 134, S. 137 Rn. 36; zu Figuren literarischer Erzählungen BGH, Urt. v. 17.7.2013 – I ZR 52/12 – *Pippi-Langstrumpf-Kostüm* = GRUR 2014, 258, S. 260 Rn. 25 ff.; vgl. weiter *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 27.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

ke in älterer Rechtsprechung mit der Begründung, dass nur die Sammlung und Strukturierung des Materials die Form darstelle, gerade aus den schutzfähigen Elementen ausgeschlossen,[323] obwohl gerade die wissenschaftlich korrekte Arbeit es erfordert, den Forschungsstand vollständig und korrekt aufzuarbeiten. Insofern könnte selbst die Materialauswahl und -sammlung bereits als den Methoden der Wissenschaft geschuldet und damit als nicht urheberschutzfähig eingeordnet werden. Weiter erscheint es unwahrscheinlich, dass die Gedankenführung ausschließlich vom Forschungsthema bestimmt wird und nicht von der Individualität des Autors geprägt sein kann.[324] Die jüngere Rechtsprechung schloss sich diesem Ausschluss insofern nicht mehr uneingeschränkt an.[325] Insgesamt wird jedoch anschaulich, welche Schwierigkeiten die Trennung von Inhalt und Form bereitet.

---

323 *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 81; *Götting* in FS Nordemann, S. 7, 8; BGH, Urt. v. 14.07.1993 - I ZR 47/91 – *Buchhaltungsprogramm* = GRUR 1994, S. 39, 39; BGH, Urt. v, 10.10.1991 - I ZR 147/89 – *Bedienungsanweisung* = GRUR 1993, S. 34, 36; BGH, Urt. v. 12.07.1990 - I ZR 16/89 – *Themenkatalog* = GRUR 1991, S. 130, 132 f.; BGH, Urt. v. 12.03.1987 - I ZR 71/85 – *Warenzeichenlexika* = GRUR 1987, S. 704, 705; BGH, Urt. v. 17.04.1986 - I ZR 213/83 – *Anwaltsschriftsatz* = GRUR 1986, S. 739, 740 f.; BGH, Urt. v. 09.05.1985 - I ZR 52/83 – *Inkasso-Programm* = GRUR 1985, S. 1041, 1074; BGH, Urt. v. 29.03.1984 - I ZR 32/82 – *Ausschreibungsunterlagen* = GRUR 1984, S. 659, 661; BGH, Urt. v. 21.11.1980 - I ZR 106/78 – *Staatsexamensarbeit* = GRUR 1981, S. 352, 353.

324 Die Individualität von Forschungstexten betont etwa *v. Moltke* in Bezug auf juristische Kommentare, *v. Moltke*, Das Urheberrecht an den Werken der Wissenschaft, S. 81 f. Gleichwohl räumt er ein, dass die individuellen Züge, die wissenschaftliche Texte aufweisen, für die Anerkennung eines Urheberschutzes zumeist nicht genügen. *Krujatz*, Open Access, S. 75 f. betont, dass die Herangehensweise stark von individuellen Fähigkeiten, Erfahrungen und Ausdrucksweise abhängt, obwohl sie möglichst sachlich formuliert sind und dies auch sein sollten; *Götting* in FS Nordemann, S. 7, 22 betont diesbezüglich gerade den größeren Spielraum für Individualität in den Geisteswissenschaften; die Gedankenführung berücksichtigt auch *Schricker* in FS GRUR, S. 1095, 1112 Rn. 43.

325 Vgl. *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 81; u. a. BGH, Urt. v. 11. 4. 2002 - I ZR 231/99 – *Technische Lieferbedingungen* = GRUR 2002, S. 958, 959: „*Bei der sprachlichen und zeichnerischen Darstellung eines technischen Regelwerks kann die urheberrechtlich geschützte Leistung in erster Linie in der Art der Sammlung, Auswahl, Einteilung und Anordnung des Stoffs liegen;* BGH, Urt. v. 16.01.1997 - I ZR 9/95 – *CB-Infobank I* = GRUR 1997, S. 459, 461: „*Soweit diese als wissenschaftliche Schriftwerke bezeichnet werden könnten, deren schöpferischer Eigentümlichkeitsgrad vornehmlich in der Form und der Art der Sammlung und Anordnung des dargebotenen Stoffes zu bemessen ist*".

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Eine weitere Eingrenzung des Schutzniveaus urheberrechtlicher Werke ergibt sich aus der Ausnahme wissenschaftlicher Erkenntnisse und Lehren aus dem Urheberrecht,[326] die zur Folge hat, dass doch gewisse inhaltliche Elemente aus dem Schutzniveau ausgenommen sind. Dogmatisch werden dafür verschiedene Begründungen herangezogen. Teilweise wird auf die Abgrenzung des Urheber- zum Patentrecht verwiesen.[327] Dieses Unterscheidungskriterium wurde mit der Begründung kritisiert, dass sich Schutzgegenstand und -richtung von Patent- und Urheberrecht voneinander unterscheiden, selbst wenn sie im Einzelfall auf denselben Gegenstand fallen sollten.[328] Teilweise wird die Ausnahme wissenschaftlicher Erkenntnisse aus dem Urheberrecht darauf gestützt, dass sie nicht erschaffen, sondern aufgefunden werden[329] sowie darauf, dass die Wissenschaft auf bestimmte Methoden angewiesen sei, mithilfe derer neue Erkenntnisse gefunden werden – aufgrund der Vorgegebenheit der Methoden fehle

326 Über diesen Grundsatz besteht in Literatur und Rechtsprechung heute weitgehende Einigkeit, obwohl er in der Gesetzesbegründung zum UrhG bewusst nicht aufgenommen wurde. BT-Drs. IV/270, S. 35. Im internationalen Recht ist er hingegen in Art. 9 Abs. 2 des TRIPS-Übereinkommens normiert, *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 40; zufolge lässt sich das als *„allgemein akzeptiertes Postulat von der urheberrechtlichen Schutzfreiheit wissenschaftlicher Entdeckungen, Erkenntnisse, Lehren und Theorien"* bezeichnen; *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 32; *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 881 f.; *Schulze* in Dreier/Schulze, UrhG, § 2 Rn. 41; schon in Bezug auf das LitUrhG BGH – Urt. v. 27.3.1963 - I b ZR 129/6 – *Rechenschieber* = GRUR 1963, S. 633; BGH – Urt. v. 21.11.1980 - I ZR 106/78 – *Staatsexamensarbeit* = GRUR 1981, S. 352, S. 353. Ursprünglich entstammt der Grundsatz, nach dem generell Ideen frei sind, der Aufklärung, *Hubmann*, UFITA 1957/II, S. 1, 1.

327 Wie etwa in BGH, Urt. v. 15. 12. 1978 - I ZR 26/77 – *Flughafenpläne* = NJW 1979, S. 1548, sowie in BGH, Urt. v. 29.03.1984 - I ZR 32/82 – *Ausschreibungsunterlagen* = GRUR 1984, S. 659; dazu eingehend *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 60.

328 V. *Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 59 ff.; ähnlich *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 112; *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 72 ff.; zum Verhältnis von Urheber- und Patentschutz *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einl. Rn. 30; BGH, Urt. v. 04.10.1990 - I ZR 139/8 – *Betriebssystem* = GRUR 1991, S. 449 m. Anm. *Betten*, S. 450.

329 *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 41; *Hubmann*, UFITA 1957/II, S. 1, 8: *„Im Bereich des Wahren gibt es also keine schützenswerte Schöpferwürde, die Jünger der Wissenschaft sind auf die reine Entdeckerehre beschränkt"*.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

auch der aufgefundenen Erkenntnis die notwendige Individualität.[330] Seine Grenzen findet diese Begründung jedenfalls im Bereich geisteswissenschaftlicher Theorien, die nicht oder nicht im gleichen Umfang methodisch bestimmt sind wie Theorien der Naturwissenschaften.[331] Das Ziel der verschiedenen Begründungen ist dasselbe: Es soll verhindert werden, dass wissenschaftlicher Fortschritt monopolisiert wird. Im Interesse der Allgemeinheit muss auf vorbestehendes Wissen ohne Verletzung von Urheberrechten zurückgegriffen werden können.[332]

Mittlerweile wurde sich von der strengen Differenzierung von Form und Inhalt weitestgehend gelöst.[333] Das ist zu befürworten, denn das Kriterium der Individualität ist bereits für sich geeignet, den urheber-

---

330 *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 86; BGH, Urt. v. 21.11.1980 - I ZR 106/78 – *Staatsexamensarbeit* = GRUR 1981, S. 352, 353; OLG Frankfurt, Urt. v. 13.06.1983 - 6 W 34/83 – *Pengo* = GRUR 1983, S. 753, 755; BGH, Urt. v. 09.05.1985 - I ZR 52/83 – *Inkasso-Programm* = GRUR 1985, S. 1041, 1047. Eine besondere Bedeutung hatte diese Thematik insbesondere, als die Frage aufkam, ob Computerprogramme urheberrechtlich geschützt sind, vgl. dazu bereits *R. Köhler,* Der urheberrechtliche Schutz der Rechenprogramme, S. 61: „*In der räumlichen Verteilung der Sinnbilder bleibt dem Programmierer nur ein äußerst keiner Spielraum*".

331 *Berking,* Die Unterscheidung von Inhalt und Form im Urheberrecht, S. 100; *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 66 f. jedenfalls auch für neue Theorien und Erkenntnisse statt für wissenschaftliche Beiträge, die vorhandene Erkenntnisse ordnen oder systematisieren; *Hubmann*, UFITA 1957/II, S. 1, 7 f.; *Plander*, UFITA 1976/I, S. 25, 41 f.

332 Im Einklang mit *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 75 wäre das Ende der freien wissenschaftlichen Auseinandersetzung zu befürchten, wenn die wissenschaftliche Erkenntnis schutzfähig wäre. Der gleichen Meinung ist *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 48. So im Grundsatz bereits *Ulmer,* Der Urheberschutz wissenschaftlicher Werke unter besonderer Berücksichtigung der Programme elektronischer Rechenanlagen, S. 15.

333 *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 43; *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 78 ff.; schon sehr frühzeitig *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 60 f.; relativierend bereits in BGH, Urt. v. 16.01.1997 - I ZR 9/95 – *CB-Infobank I* = GRUR 1997, S. 459, 461 f. eindeutiger dann in BGH, Urt. v. 11. 4. 2002 - I ZR 231/99 – *Technische Lieferbedingungen* = GRUR 2002, S. 958, 959; LG Köln, Urt. v. 14.01.1972 - 5 O 401/71 – *Kinder in Not* = GRUR 1973, S. 88, 281; BGH, Urt. v. 1. 12. 2010 - I ZR 12/08 – *Perlentaucher* = GRUR 2011, S. 134, 137 Rn. 36; KG, Beschl. v. 12.03.2014 – 24 W 21/14 – *Kein Urheberrechtsschutz für Urteilsanalyse* = ZUM 2014, S. 969, 969 f.; vgl. auch die Nachweise in Fn. 319.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

rechtlichen Schutzumfang angemessen zu bestimmen[334] und ermöglicht zugleich eine flexiblere Beurteilung. Problem- bzw. Grenzfälle wie etwa die fehlende Schutzfähigkeit geisteswissenschaftlicher Theorien, die dennoch von der Persönlichkeit desjenigen geprägt sein können,[335] der sie begründete, lassen sich gleichwohl mittels des ungeschriebenen Grundsatzes lösen, dass Wissen als solches aus dem Urheberrechtsschutz ausgenommen ist.[336] Auch der EuGH betonte, dass trotz der weitgehend abgesenkten Schutzvoraussetzungen für die Anerkennung der für einen Werkschutz notwendigen Originalität bzw. der „*eigenen geistigen Schöpfung*" überhaupt ein Gestaltungsspielraum erforderlich ist, in dem der Urheber seine Persönlichkeit ausdrückt.[337] Im Bereich wissenschaftlichen Werkschaffens beschränkt sich dieser Gestaltungsspielraum wiederum auf die anschauliche Erläuterung, Darstellung oder sonstige Vermittlung wissenschaftlicher Erkenntnisse.[338] Auch dann bleibt die sog. „*gewöhnliche wissenschaftliche*

334 So schon *Ulmer,* Der Urheberschutz wissenschaftlicher Werke unter besonderer Berücksichtigung der Programme elektronischer Rechenanlagen, S. 13 ff.; andeutungsweise auch *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 48, der es als konsequenter einordnet; ebenso *Waiblinger,* "Plagiat" in der Wissenschaft, S. 73; im Grunde auch schon *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 81 f.; ebenso *R. Köhler,* Der urheberrechtliche Schutz der Rechenprogramme, S. 53.

335 Anderer Auffassung war noch *Ulmer,* Der Urheberschutz wissenschaftlicher Werke unter besonderer Berücksichtigung der Programme elektronischer Rechenanlagen, S. 13 ff., der gar von einer Austauschbarkeit der Wissenschaftler ausging.

336 So auch bereits *Haberstumpf,* Zur Individualität wissenschaftlicher Sprachwerke, S. 68; 81 f.

337 Technische Gegebenheiten können den Gestaltungspielraum daher bis auf Null herabsetzen, *Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 6; EuGH, ECLI:EU:C:2011:631 – *Football Association Premier League u. Murphy* = GRUR 2012, S. 156, 160 Rn. 96; EuGH, ECLI:EU:C:2010:816 – *BSA/Kulturministerium* = GRUR 2011, S. 220, 222 Rn. 96; EuGH, ECLI:EU:C:2011:798 – *Painer/Standard* = GRUR 2012, S. 166, 168 Rn. 87 ff. EuGH, ECLI:EU:C:2010:816 – *BSA/Kulturministerium* = GRUR 2011, S. 220, 222 Rn. 96 enthält explizit, dass das Kriterium der Originalität nicht dann erfüllt sein kann, wenn der Ausdruck durch die technischen Komponenten vorgegeben ist, „*denn die verschiedenen Möglichkeiten der Umsetzung einer Idee sind so beschränkt, dass Idee und Ausdruck zusammenfallen*"; darauf verweisend EuGH, ECLI:EU:C:2012:115 – *Football Dacato/Yahoo* = GRUR 2012, S. 386, 338 Rn. 38 f.; ausführlich zu der Rechtsprechung des EuGH *E.-M. König,* Der Werkbegriff in Europa, S. 26 ff.

338 *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 80. Dabei können wissenschaftliche Werke auch, wenn auch nur in begrenztem Umfang, ästhetische Gefühle wecken. Rationalität schließt Ästhetik also nicht grundsätzlich aus, vgl. *v. Moltke,* Das Urheberrecht an den Werken der Wissen-

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Ausdrucksweise", also die Fachsprache, regelmäßig schutzlos.[339] Das führt letztlich zu dem Ergebnis, dass je wissenschaftlicher ein wissenschaftlicher Text formuliert und aufgebaut ist, desto niedriger auch das Schutzniveau einzuordnen ist.

Auch an der Gewährung des auf der Datenbank-RL basierenden Datenbankherstellerrechts wird kritisiert, dass es dazu führe, dass ursprünglich urheberrechtlich freie Inhalte monopolisiert würden.[340] Diesem Vorwurf ist differenziert zu begegnen. Einerseits schützt das Datenbankherstellerrecht nicht die enthaltenen Daten selbst, sondern ihre Zusammenstellung, Art. 7 Abs. 4 S. 2 Datenbank-RL, d. h. es ist weder die Datenerzeugung noch der Ausschluss Dritter von Nutzung unwesentlicher Datenbankteile erfasst.[341] Andererseits bewirkt die Einschränkung des Art. 7 Abs. 1 Datenbank-RL, dass eine Datenbank nur dann geschützt wird, wenn ihre

---

schaft, S. 26; *Haberstumpf*, Zur Individualität wissenschaftlicher Sprachwerke, S. 24 betont die Beschränktheit des Gestaltungsspielraums bzw. die Abhängigkeit der Wissenschaftssprache vom jeweiligen Fachgebiet, räumt aber ein, dass es auch anschauliche wissenschaftliche Erzeugnisse gibt, die in ihrer Darstellungweise ausreichend individuell sind.

339 Das Besagte gilt insbesondere für besonders sachlich gehaltene Ausführungen; *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 86 weist darauf hin, dass die übliche Darstellungsweise gerade durch die Fachrichtung vorgegeben ist und insofern kein Raum für eine schöpferische Gedankenführung bleibe; ebenso *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 93; nach *v. Moltke*, Das Urheberrecht an den Werken der Wissenschaft, S. 78, 185 f. gilt dies für „*allgemein üblichem standardisierte Kunstsprachen und Darstellungsmethoden*"; ähnlich *Kuschel*, OdW 2020, S. 43, 44 f. vgl. die entsprechende Rechtsprechung, u. a. BGH, Urt. v. 21.11.1980 - I ZR 106/78 – *Staatsexamensarbeit* = GRUR 1981, S. 352, 355; BGH, Urt. v. 29.03.1984 - I ZR 32/82 – *Ausschreibungsunterlagen* = GRUR 1984, 659, S. 663; ursprünglich und (zu) weitgehend *Ulmer*, Der Urheberschutz wissenschaftlicher Werke unter besonderer Berücksichtigung der Programme elektronischer Rechenanlagen, S. 13 ff.

340 *Peukert*, JIPITEC 2013, S. 142, 143 Rn. 5 f. zufolge verschafft das digitale Urheberrecht auf diese Weise „*die volle Herrschaft über wissenschaftliche Information*"; *Flechsig*, ZUM 1997, S. 577, 590; *Ullmann* in FS Brandner, S. 507, 523 f.

341 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 87a Rn. 13; *Dietz* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 87a Rn. 49; *Hermes* in Wandtke/Bullinger, UrhG, 5. Aufl. 2019, § 87a Rn. 36; EuGH, ECLI:EU:C:2004:696 – *Fixtures-Fußballspielpläne I* = GRUR 2005, S. 252, 256 Rn. 40.; EuGH, ECLI:EU:C:2012:115 – *Football Dacato/Yahoo* = GRUR 2012, S. 386, 387 Rn. 32 ff.; EuGH, ECLI:EU:C:2004:695 – *British Horseracing Board* = GRUR 2005, S. 244, 247 Rn. 34; so bereits zuvor *Leistner*, Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht, S. 150 f.; dazu umfassend *Rieger*, Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 134 ff. Zu den Details des Abgrenzungskriteriums ausführlich *Leistner*, CR 2018, S. 17 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Beschaffung, Überprüfung oder die Darstellung ihres Inhalts eine nach Art oder Umfang wesentliche Investition erforderte, aufgrund ihrer Konturlosigkeit[342] eine erhebliche Einschränkung, denn die Wesentlichkeit kann sowohl nach Art als auch nach Umfang bestimmt werden.[343] Insofern ist das Datenbankherstellerrecht auch dann betroffen, wenn nur quantitativ kleine Bestandteile extrahiert werden, die aber qualitativ wesentlich sind.[344] Das ist nach allgemeiner Auffassung jedenfalls der Fall, wenn für diese Bestandteile der Einsatz wesentlicher Mittel erforderlich war, denn dann wird die geschützte Investition am stärksten gefährdet.[345] Das erschwert es erheblich einzuschätzen, in welchem Umfang Daten aus Datenbanken extrahiert werden dürfen. Daneben ist der Kritik am Datenbankherstellerrecht dahingehend, dass die in der Datenbank enthaltenen unabhängigen Elemente gerade nicht in jedem Fall selbst zu einer solchen Datenbank zusammengeführt werden können, wenn auf die Daten schon nicht zugegriffen werden kann,[346] zuzustimmen. Insofern schützt das Datenbankherstellerrecht mittelbar doch die Erzeugung der Daten bzw. die Daten selbst.[347] Dennoch muss eingeräumt werden, dass auch unabhängig des Leistungsschutzrechts kein faktischer Datenzugang bestünde, sofern die Datenbank nicht aus allgemein verfügbaren Daten besteht.[348]

---

342 Ähnlich *Vogel*, ZUM 1997, S. 592, 596.

343 So Art. 7 Abs. 1 Datenbank-RL.

344 *Truyens/van Eecke*, CLSR 2014, S. 153, 161.

345 *Vohwinkel* in BeckOK UrhG, 33. Edition 2022, § 87b Rn. 13; *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 87b Rn. 7; *Czychowski* in Fromm/Nordemann, UrhG, § 87b Rn. 6 ff.; *Truyens/van Eecke*, CLSR 2014, S. 153, 161; *Lehmann*, CR 2005, S. 15, 16; *Benecke*, CR 2004, S. 608, 612 f.; EuGH, ECLI:EU:C:2004:695 – *British Horseracing Board* = GRUR 2005, S. 244, 250 Rn. 69 ff.

346 *Leistner*, Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht, S. 233 ff. Das Copyright Office der Vereinigten Staaten sieht ebenfalls Datenbanken, in denen der Datenbankhersteller gleichzeitig der Datenerzeuger ist, kritisch, U. S. Copyright Office, Report on legal protection for databases (August 1997), S. 101 ff.

347 So u. a. *Leistner*, Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht, S. 233, das soll auch der Data Act klarstellen, Europäische Kommission, Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on dair access to and use of data (Data Act), COM(2022) 68 final (23.02.2022)Art. 35 sowie Erw. 84 , vgl. dazu näher auf S. 255 f.

348 Dennoch wurde bei Einführung des Datenbankherstellerrechts gefordert, die Datenbanken einer Zwangslizenz zu unterwerfen, wenn die enthaltenen Daten nicht unabhängig von der Datenbank, d. h. mit anderen Mitteln beschafft werden können, um einer Informationsmonopolisierung entgegenzuwirken, vgl. Art. 11 des RichtlinienentwurfsKOM (93) 464 endg. - SYN 393 vom 04.10.1993;

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

2. Die urheberrechtliche Ein- und Zuordnung von Forschungsdaten

Anhand von Forschungsdaten, die in jüngerer Zeit stärker in das wissenschaftliche Interesse geraten sind,[349] werden die dargelegten Grundsätze, die den Schutz wissenschaftlicher Erzeugnisse betreffen, anschaulich. Zu unterscheiden sind der Schutz der Forschungsdaten selbst, der im Einzelnen spezifiziert werden muss,[350] und der Schutz ihrer Sammlung. Nicht thematisiert wird die generelle Frage der Zuordnung von Daten.[351]

---

bedauernd *Ullmann* in FS Brandner, S. 507, 524; als angemessen erachtete diese Vorgabe auch *v. Gamm*, GRUR 1993, S. 203, 204; ebenso befürwortet *Hoebbel*, CR 1993, S. 12, 17 eine entsprechende Regelung.

349 Besonderes Interesse wecken Forschungsdaten, weil sich ihre Sammlung und Aufbereitung zumeist aufwendig gestaltet und daher ein grundsätzliches Interesse daran besteht, weiteren (wissenschaftlichen oder wirtschaftlichen) Nutzen aus ihnen zu ziehen. Das Interesse an Forschungsdaten und der effektiven Nutzung beweisen etwa Forschungsprojekte, die sich gerade mit diesen Thematiken auseinandersetzen, genannt seien u. a. das NFDI-Konsortium, ein Projekt der Gemeinsamen Wissenschaftskonferenz, näher dazu unter https://www.nfdi.de/, das BMBF-geförderte DataJus-Projekt, https://tu-dresden.de/gsw/phil/irget/jfbimd13/forschung/forschungsprojekt-datajus, sowie XSample, ein Projekt des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg, https://www.izus.uni-stuttgart.de/fokus/fdm-projekte/xsample/. Auch die Kultusministerkonferenz hat schon mehrfach gefordert, einen ausreichenden Rechtsrahmen zur Weiterverwendung von Forschungsdaten zu bilden, vgl. Kultusministerkonferenz, Referentenentwurf des Bundesministeriums der Justiz und für Verbraucherschutz (BMJV) für ein Gesetz zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarkts; Stellungnahme der Kultusministerkonferenz (05.11.2020), S. 4. Dem schloss sich jüngst auch der Bundesrat an, vgl. BR-Drs. 142/21 (Beschluss), S. 12 f. Vgl. in Bezug auf das Interesse der Forschung, die TDM-Korpora nachzunutzen, die späteren Ausführungen auf S. 259 ff.

350 Ebenso *Kuschel*, OdW 2020, S. 43, 44 mit Bezug auf BGH, Urt. v. 27.09.1990 - I ZR 244/88 – *Grabungsmaterialien* = GRUR 1991, S. 523, 525.

351 In Bezug auf die Zuordnung von Daten wird nicht nur datenschutzrechtlich und urheberrechtlich gesprochen, auch eine zivilrechtliche Zuordnung wurde mittels verschiedener Rechtskonstrukte schlagwortartig vorgeschlagen bzw. untersucht. Die Diskussion um ein Dateneigentum ist aufgrund von Problematiken mit der Vereinbarkeit mit der informationellen Selbstbestimmung in Bezug auf eine Ökonomisierung von Daten und auch abseits dessen aufgrund fehlender Praktikabilität weitestgehend abgeflaut, vgl. dazu u. a. grundlegend *Zech*, Information als Schutzgegenstand; *Hoeren*, MMR 2013, S. 486 ff.; befürwortend in Bezug auf verhaltensgenerierte Personendaten *Fezer*, MMR 2017, S. 3, 5 ff.; *Fezer*, ZGE 2017, S. 356, 356; *Fezer*, ZD 2017, S. 99 ff.; kritisch *Determann*, MMR 2018, S. 277, 277 f.; *Determann*, ZD 2018, S. 503, 508; *Specht*, ZGE 2017, S. 411, 414 f.; *Spindler*, ZGE 2017, S. 399, 400 ff.; *Kühling/Sackmann*, ZD 2020, S. 24 ff.;

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Grundsätzlich kommt in Bezug auf die Forschungsdaten selbst ein Schutz als Darstellungen wissenschaftlicher oder technischer Art, §§ 2 Abs. 1 Nr. 5, Abs. 2 UrhG, oder als unbenannte Werke nach § 2 Abs. 2 UrhG in Betracht.[352] Für einen urheberrechtlichen Werkschutz ist es grundsätzlich nicht schädlich, wenn – wie es bei Forschungsdaten der Fall ist – eine wissenschaftliche oder eine technische Zweckrichtung verfolgt wird.[353] Entsprechend der vorangegangenen Erwägungen kommt ein urheberrechtlicher Werkschutz gerade bei der Darstellung von Fakten allerdings nur dann in Betracht, wenn sie besonders individuell gestaltet ist. Bei einer möglichst sparsamen und analysefreundlichen Darstellung ist für Individualität zumeist kein Raum, denn Ziel ist gerade die neutrale Abbildung bestimmter Zustände. Handelt es sich bei Forschungsdaten um Messwerte, Rohdaten, tabellarische Anordnungen, die sich auf Zustände in der Natur, auf andere wissenschaftliche Studien oder auf Erhebungen über die Realität beziehen, sind diese infolgedessen nur in den seltensten Fällen urheberrechtlich geschützt.

Anders verhält es sich, wenn die Forschungsdaten auf Kulturdaten oder Texten basieren: Sind jedenfalls Teile von Texten oder sonstigen Medien, die Sprache enthalten, Musikstücken, Lichtbildwerke in den Forschungsdaten enthalten, sind diejenigen Forschungsdaten solange urheberrechtlich geschützt, wie sie (noch) die persönlichen geistigen Schöpfungen der Ursprungswerke enthalten.[354] Daneben kommen auch diverse Leistungsschutzrechte in Betracht, genannt seien etwa Lichtbilder (§ 72 UrhG), Filme (§ 95 UrhG), Tonträger (§ 85 UrhG) oder wissenschaftliche Ausgaben

---

Stender-Vorwachs/Steege, NJOZ 2018, S. 1361 ff.; *Peukert* in FS Schricker, S. 149, 152 ff.; nicht endgültig *Wiebe/Schur*, ZUM 2017, S. 461, 472 f. Außerdem kam eine Diskussion um die Notwendig- bzw. Sinnhaftigkeit eines Datenbesitzes auf (befürwortend etwa *Hoeren*, MMR 2019, S. 5 ff., 7 f.; *Michl*, NJW 2019, S. 2729 ff.; ablehnend *Martini/Kolain et al.*, MMR-Beilage 2021, S. 3, 13; *Zech*, Information als Schutzgegenstand, S. 347.). Inzwischen wird auch eine Datentreuhand diskutiert, *Specht-Riemenschneider/Blankertz*, MMR 2021, S. 369370 f.; *Specht-Riemenschneider/Blankertz et al.*, MMR-Beilage 2021, S. 25 ff. Die „*faktische Macht, über die Integrität eines digitalen Datums bestimmen zu können*" betrachten *Martini/Kolain et al.*, MMR-Beilage 2021, S. 3, 16 als „*Datenhoheit*".

352 Das hat der BGH in BGH, Urt. v. 27.09.1990 - I ZR 244/88 – *Grabungsmaterialien* = GRUR 1991, S. 523, 525 in Erwägung gezogen.

353 *Kuschel*, OdW 2020, S. 43, 44; BGH 10.5.1984 – I ZR 85/82 – *Elektrodenfabrik* = GRUR 1985, 129, 130

354 Dazu wird neuerdings das Kriterium der Wiedererkennbarkeit angelegt, s. dazu an späterer Stelle S. 113 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

(§ 70 UrhG).[355] Die benannten Werktypen können insofern in Sammlungen von Forschungsdaten auftreten bzw. Forschungsdatensätze können ein „*Konglomerat*" an verschiedenen Schutzrechten neben nicht geschützten Inhalten enthalten.[356] Sofern diese Schutzgegenstände so verarbeitet wurden, dass sie nur noch die urheberrechtlich ungeschützten Informationen enthalten, d. h. deren urheberrechtlich geschützte Darstellungen von den beinhalteten Informationen gelöst wurden, sind sie gemeinfrei.[357] Aus pragmatischen Gründen muss in der Praxis jedoch im Zweifelsfall von einem Urheberrechtschutz ausgegangen werden.[358]

Auch die Zusammenstellung der Forschungsdaten kann schutzfähig sein. Dieser Schutz kann sich aus dem Schutz von Sammelwerken, §§ 4 Abs. 1, 2 Abs. 1 UrhG, dem Datenbankwerkschutz, §§ 4 Abs. 2, 2 Abs. 2 UrhG oder aus dem Datenbankherstellerrecht, §§ 87a ff. UrhG ergeben. Sammelwerke schützen Sammlungen von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund der Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung darstellen. Der Schutz von Datenbankwerken, §§ 4 Abs. 2, 2 Abs. 2 UrhG, erfordert dagegen, dass die Auswahl oder Anordnung von voneinander unabhängigen Elementen eine persönliche geistige Schöpfung ist und deren Elemente systematisch oder methodisch angeordnet sind und einzeln zugänglich sind. Es handelt sich also gerade nicht um einheitliche Werke.[359] In beiden Fällen muss sich die Originalität bzw. die Individualität in der Auswahl oder Anordnung des Inhalts zeigen.[360] Aus einem „*Mehr*" an Wissenschaftlichkeit ergibt sich insofern nicht nur in Bezug auf wissenschaftliche Textpublikationen oder anderweitige Darstellungen von Forschungsergeb-

355 *Lauber-Rönsberg* in: Putnings/Neuroth/Neumann (Hrsg.), Praxishandbuch Forschungsdatenmanagement, S. 89, 90.

356 *Lauber-Rönsberg* in: Putnings/Neuroth/Neumann (Hrsg.), Praxishandbuch Forschungsdatenmanagement, S. 89, 91.

357 *Rack/F. Boehm/Pasdzierny/D. Schmidt* in: Schrör/Fischer/Beaucamp/Hondros (Hrsg.), Tipping Points, S. 253, 268 nennen hierzu etwa Harmonien, die aus Musikwerken extrahiert werden. Ähnlich ist der von *de la Durantaye/B. Raue*, RuZ 2020, S. 83 ff.; *Erler*, RuZ 2020, S. 108 ff.; *Grisse*, RuZ 2020, S. 143 ff.; *Jotzo*, RuZ 2020, S. 128 ff.; *B. Raue/Schöch*, RuZ 2020, S. 118 ff.; *Schöch/Döhl et al.*, RuZ 2020, S. 160 ff. dargestellte Ansatz.

358 *Rack/F. Boehm/Pasdzierny/D. Schmidt* in: Schrör/Fischer/Beaucamp/Hondros (Hrsg.), Tipping Points, S. 253, 268.

359 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 4 Rn. 1.

360 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 4 Rn. 11, 19; *Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 4 Rn. 19.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

nissen ein „*Weniger*" im Schutzniveau. Das gilt insbesondere bei einer logischen Anordnung von Messdaten.[361]

Datenbanken, die keine urheberrechtlichen Werke sind, sind gem. § 87a Abs. 1 UrhG nur dann geschützt, wenn die Beschaffung, Überprüfung oder Darstellung der Daten oder anderen unabhängigen Elementen eine nach Art oder Umfang wesentliche Investition erfordert.

Insofern muss bei der Beantwortung der Frage nach der Schutzfähigkeit von Forschungsdaten differenziert werden. Schutzfähig sind einzelne Daten in der Regel dann, wenn sie geschützten Gegenständen entnommen wurden. Die Sammlung von Forschungsdaten kann sowohl als Sammelwerk und Datenbankwerk (§ 4 Abs. 1, 2, § 2 Abs. 2 UrhG) geschützt sein als auch als Datenbank gem. §§ 87a ff. UrhG geschützt sein.

II. Die besonderen Interessen im Wissenschaftsurheberrecht

Ausgehend von der Feststellung, dass das Schutzniveau wissenschaftlicher Werke im Gegensatz zum sonstigen Urheberrecht eher niedrig angesiedelt ist, ist es notwendig, die Interessenlage im Wissenschaftsurheberrecht zu untersuchen, um zu verstehen, wie die Wissenschaft durch das Urheberrecht gesteuert wird bzw. werden kann. Dass ein „*Mehr*" an Wissenschaftlichkeit ein „*Weniger*" an urheberrechtlicher Schutzhöhe zur Folge hat, erstaunt[362] und wird in der Literatur auch teilweise kritisch betrachtet, weil *„negative Auswirkungen auf das wissenschaftliche Publizieren und die Kommunikationskultur innerhalb der Wissenschaftsgemeinde"* befürchtet werden.[363] Das, was die Wissenschaft zunächst zu benachteiligen scheint, soll ihr jedoch gerade dienen, indem das in den wissenschaftlichen Werken enthaltene Wissen möglichst frei kursieren kann.[364]

Die Interessenlage im Bereich wissenschaftlicher Urheber unterscheidet sich allerdings auch erheblich von der Interessenlage der Urheber, die der Vorstellung des historischen Gesetzgebers zugrundeliegt, der einerseits die

361 *Kuschel*, OdW 2020, S. 43, 44; *T. Hartmann* in: Taeger (Hrsg.), Law as a service (LaaS), DSRITB 2013, S. 505, 512; *T. Hartmann*, InTeR 2013, S. 199, 202.

362 *Geiger/Frosio/Bulanyenko*, IIC 2018, S. 814, 819 f.

363 *Haberstumpf*, ZUM 2012, S. 529, 530 warf dem zugrundeliegenden Urteil, OLG Frankfurt, Urt. v. 27.03.2012 – 11 U 66/11 = ZUM 2012, S. 574, sogar *„wissenschaftsfeindliche Tendenzen"* vor; *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 44 sieht das urheberrechtliche Schutzniveau ebenso kritisch.

364 *Kuschel*, OdW 2020, S. 43, 43.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

wirtschaftliche Existenz des Urhebers und andererseits die persönlichkeitsrechtlichen Interessen sichern wollte, § 11 UrhG.[365] Besonders relevant für wissenschaftliche Urheber ist das Urheberpersönlichkeitsrecht, insbesondere das Veröffentlichungsrecht, § 12 UrhG, das Recht des Schutzes vor Entstellung seines Werks, § 14 UrhG und insbesondere das Interesse, als Urheber anerkannt zu werden, § 13 UrhG,[366] denn diese Bekanntheit bestimmt die wissenschaftliche Reputation und die Karriereaussichten.[367] Damit geht auch ein Interesse an einer möglichst großen Verbreitung ihrer Erkenntnisse einher.[368] Die wirtschaftlichen Interessen treten demgegenüber eher in den Hintergrund, denn wissenschaftliche Urheber finanzieren sich zumeist als Arbeitnehmerurheber, d. h. durch ihr Gehalt aus ihrer Tätigkeit im Wissenschaftsbetrieb.[369] Die Verwertungsinteressen liegen dabei hauptsächlich bei den Wissenschaftsverlagen.[370] Der Primärmarkt wissenschaftlicher Publikationen und damit auch der Wissenschaftsverlage wird zu einem großen Anteil von wissenschaftlichen Bibliotheken bedient.[371] Ob es eine finanzielle Beteiligung der Urheber gibt, ist je nach Fachdis-

365 *Loewenheim/Pfeifer* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 11 Rn. 1; BT-Drs. IV/270, S. 43.

366 *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 171; *Haedicke* in: Leible/Ohly/Zech (Hrsg.), Wissen - Märkte - Geistiges Eigentum, S. 93, 101 f.; *Pfeifer*, GRUR 2009, S. 22 ff., 23; *Taubert/Weingart* in: Sutter/Mehler (Hrsg.), Medienwandel als Wandel von Interaktionsformen, S. 159, 169; *Leistner/Hansen*, GRUR 2008, S. 479, 486; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 38 Rn. 25; *Peukert*, JIPITEC 2013, S. 142, 135 Rn. 20; *Suber,* Open access, S. 10: „*The academic custom to write research articles for impact rather than money may be a lucky accident that could have been otherwise. Or it may be a wise adaptation that would eventually evolve in any culture with a serious research subculture.*".

367 *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 171.

368 *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 171; BGH, Urt. v. 5. 7. 2001 - I ZR 311/98 – *Spiegel CD-ROM* = GRUR 2002, 248, S. 251.

369 *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 171; *Krujatz,* Open Access, S. 26; *Peukert*, JIPITEC 2013, S. 142, 145 Rn. 20; *Kuhlen*, ZGE 2015, S. 77, 116; *P. Katzenberger*, Die Naturwissenschaften 1975, S. 555, 557. Zwar ist z. B. im juristischen Bereich auch die Publikation von Lehrbüchern mit großer Auflage oder Standard-Kommentaren lukrativ, dies betrifft aber nur einen sehr geringen Prozentsatz der wissenschaftlichen Autoren.

370 *Haucap/Loebert/Spindler/Thorwarth,* Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, S. 99.

371 *Haucap/Loebert/Spindler/Thorwarth,* Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, S. 101; *Herrmann*, ZUM 2016, S. 496, 496.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

ziplin unterschiedlich und kommt zudem auf die Art der Publikation an.[372] Die Abhängigkeit wissenschaftlicher Autoren von Publikationen ist insofern zumeist nicht wirtschaftlicher Natur, sondern notwendig zur Erlangung wissenschaftlichen Renommees.[373] Die Erträge aus wissenschaftlichen Publikationen dienen hierbei zumeist als Nebeneinnahme und sind nur in den seltensten Fällen echter Gehaltsbestandteil. Zwar kann sich eine große Anzahl relevanter Publikationen langfristig auch auf das Gehalt auswirken.[374] Relevant ist zu diesem Zweck aber dennoch, dass wissenschaftliche Urheber auch dann, wenn nur wenige Bestandteile ihrer Erzeugnisse die notwendige Schöpfungshöhe erreichen, für das Gesamtwerk die Anerkennung als wissenschaftliche Urheber erlangen. Entscheidend ist insofern erneut das durch die Veröffentlichungen erlangte Renommee.

Es zeigt sich, dass im Bereich wissenschaftlicher Urheber die persönlichkeitsrechtsbasierten Bestandteile viel eher ins Gewicht fallen als die verwertungsrechtlichen. Das beweist auch der Erfolg offener Publikationsformate, indem der sog. grüne Weg, d. h. ein Zweitveröffentlichungsrecht nach Ablauf einer einjährigen Frist[375] in § 38 Abs. 4 UrhG verankert worden ist.[376] Problematisch ist in einigen Disziplinen (noch) die wissenschaftli-

372 Gerade Monografien werden i. d. R. nur unter der Bedingung eines Honorarverzichts oder bzw. sowie der Zahlung eines Druckkostenzuschusses abhängig gemacht, *Haucap/Loebert/Spindler/Thorwarth,* Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, S. 99.

373 Etwa *Krujatz,* Open Access, S. 24; *Euler*, RuZ 2020, S. 56, 63 verweisen in diesem Kontext auf die Phrase *„publish or perish"*, was übersetzt etwa *„veröffentliche, oder gehe unter"* bedeutet. Diese verdeutlicht den Publikationsdruck, unter dem wissenschaftliche Urheber stehen. Ebenso *Haedicke* in: Leible/Ohly/Zech (Hrsg.), Wissen - Märkte - Geistiges Eigentum, S. 93, 100.

374 *Peukert*, JIPITEC 2013, S. 142, 145 Rn. 20; zum Selbstzweck des wissenschaftlichen Arbeitens aus systemtheoretischer Perspektive *Luhmann,* Die Wissenschaft der Gesellschaft, S. 446 f.; ausführlich *Krujatz,* Open Access, S. 172; *Haucap/Loebert/Spindler/Thorwarth,* Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, S. 100; u. a. *Tuckmann/Leahey*, Journal of Political Economy 1975, S. 951 ff.

375 Der goldene Weg meint hingegen eine offene und kostenfreie Erstveröffentlichung, vgl. dazu ausführlich u. a. *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 180; *Krujatz,* Open Access, S. 57, 136 ff.; *Bruch/Pflüger*, ZUM 2014, S. 389, 389 f.

376 Nach § 38 Abs. 4 UrhG haben Urheber wissenschaftlicher Beiträge, die in periodisch erscheinenden Formaten wie etwa Zeitschriften erscheinen, ein unabdingbares Zweitveröffentlichungsrecht nach Ablauf einer einjährigen Embargofrist, sofern die betreffenden Beiträge im Rahmen von mindestens zur Hälfte mit öffentlichen Mitteln finanzierten Forschungstätigkeit entstanden sind, *Bruch/Pflüger*, ZUM 2014, S. 389 ff. Mit dieser urhebervertragsrechtlichen

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

che Reputation, die teilweise noch von bestimmten renommierten Fachmagazinen abhängt und deren Einbußen nicht zugunsten eines offenen Zugangs in Kauf genommen werden können.[377] Dies gilt insbesondere für die Rechtswissenschaft.[378] Doch auch in Wissenschaftsdisziplinen, in denen offene Publikationen noch nicht selbstverständlich sind, mehren sich offene Publikationsformate oder es sind jedenfalls erste Schritte erkennbar, die den Weg für offenen Zugang ebnen könnten. Es entwickeln sich dabei auch neue Finanzierungsmöglichkeiten für die kostspielige redaktionelle Arbeit an Wissenschaftsliteratur.[379]

Ein deutlicher Vorteil der Publikation in digitalen und offenen Formaten ist die größere Reichweite, die das Wissen dadurch erfährt.[380] Außerdem erhalten wissenschaftliche Urheber die notwendige Reputation umso mehr, während sie von der kommerziellen Verwertung zumeist ohnehin nicht profitieren. Damit erfüllen die offenen Publikationsmöglichkeiten ihre wesentlichen Interessen an ihren eigenen Werken.

Nicht ausführlich thematisiert wurde bislang die Kehrseite des Schutzes wissenschaftlicher Erzeugnisse und der fehlenden Notwendigkeit nach einem höheren Schutzniveau, d. h. die Erlaubnis der Wissenschaft, auf fremden Werken aufzubauen und diese in urheberrechtlich relevantem

---

Regelung wurde die Problematik umgangen, dass das EU-Recht bislang keine Erleichterungen für offene Publikationswege vorsieht.

377 *Haedicke* in: Leible/Ohly/Zech (Hrsg.), Wissen - Märkte - Geistiges Eigentum, S. 93, 100.

378 *Euler*, RuZ 2020, S. 56, 77 f. hat hierfür Publikationszahlen (Stand: Januar 2020) in den Rechtswissenschaften analysiert und ist zum Ergebnis gekommen, dass offene Publikationsformate international zu diesem Zeitpunkt lediglich 1 Prozent der Publikationen im Zeitschriftenbereich ausmachten. In der „*strukturkonservativen*" Rechtswissenschaft bestünden in Bezug auf offene Publikationsformate ein Misstrauen gegenüber neuen Publikationsformaten, so *Duve*, vgl. den Bericht von *Hellmund*, RuZ 2021, S. 164, 164 f.

379 Die RuZ stellt in diesem Zuge das erste Open Access Magazin dar, das deutschsprachig in einem renommierten Fachverlag erschienen ist und beschäftigt sich gerade mit den Themen des freien Zugangs, ist insofern gewissermaßen eine Verkörperung seiner Ideologie. Das Finanzierungsmodell ist dabei gänzlich neu, vgl. dazu *Euler*, RuZ 2020, S. 56, 81 f.

380 *Lutz*, Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 172; *Wildgans*, ZUM 2019, S. 21, 27; *Peukert*, JIPITEC 2013, S. 142, 115 Rn. 13; ein durchwachsenes Bild zeichnen allerdings noch *Mueller-Langer/Scheufen* in: Handke/Towse (Hrsg.), Handbook of the Digital Creatice Economy, S. 365, 369 f.; *Fröhlich*, IWP 2009, S. 253, 255; zum amerikanischen Markt und offen publizierten Zeitschriften der juristischen Fakultäten *Watson/Donocan/Osborne*, Edison L. & Tech. 2015, 8, die Vorteile in der Reichweite auf 15 Jahre gerechnet von über 50 Prozent feststellen.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Umfang zu nutzen sowie das diesbezüglich gesteigerte Bedürfnis. Wie bereits erläutert, stellt die Wissenschaft ihre Geisteserzeugnisse der Allgemeinheit auf eine Art und Weise zur Verfügung, aufgrund derer zwar zumeist nicht die Gesamtwerke, aber jedenfalls viele wesentliche Elemente ihrer Werke ungeschützt bleiben. Zu diesem Zwecke sind die Vertreter der Wissenschaft jedoch in gesteigertem Maße davon abhängig, auf Vorbestehendem aufzubauen, d. h. einerseits andere wissenschaftliche und andererseits auch nicht-wissenschaftliche Schutzgegenstände nutzen zu können.[381] Insofern können Schrankenbestimmungen ein wichtiges Gegengewicht zum niedrig angesiedelten Schutzniveau und einen legitimen Ausgleich bilden.[382]

Im Rahmen dieser Überlegungen darf nicht der Eindruck entstehen, das Wissenschaftsurheberrecht thematisiere die Interessen einiger weniger,[383] denn die Interessen der Wissenschaft sind die Interessen der Allgemeinheit, d. h. sämtliche wissenschaftlichen Eigengesetzlichkeiten, entsprechen den Interessen der Allgemeinheit an wissenschaftlichem Fortschritt. Das wird nicht nur daran sichtbar, dass gerade die Sozialpflichtigkeit des Eigentums Grund dafür ist, dass der Gesetzgeber dessen Inhalt erst durch Gesetz bestimmt, sondern auch, dass das Wissen als solches aus dem Bereich des Urheberrechts ausgenommen ist.[384]

---

381 *Haedicke* in: Leible/Ohly/Zech (Hrsg.), Wissen - Märkte - Geistiges Eigentum, S. 93, 100 bezeichnet das als „*single source situation*".

382 Vgl. dazu speziell in Bezug auf das TDM die Ausführungen im zweiten Teil der Arbeit, insbesondere auf S. 191 ff.

383 Sehr deutlich formuliert das *Kuhlen,* Die Transformation der Informationsmärkte in Richtung Nutzungsfreiheit, S. 23: „*Wissenschaftliche Interessen sind keine Partikularinteressen. Ein wissenschaftsfreundliches Urheberrecht ist nicht bloß eine Forderung aus der Binnensicht der Wissenschaft, sondern rechtfertigt sich durch den Anspruch der Öffentlichkeit auf freien Zugang und freie Nutzung des in der Wissenschaft produzierten und dann auch in Werken publizierten Wissens. Für die Öffentlichkeit ist das aber nicht nur Anspruch, sondern auch die Verpflichtung, den wissenschaftlich Arbeitenden die für ihre Arbeit nötigen Ressourcen zu sichern. Zu diesen Ressourcen gehören auch die publizierten Arbeiten Anderer. Ohne diese aus dem Gemeinwohlinteresse begründete Sicherung der Ressourcen könnte sich Wissenschaft nicht entwickeln. Die freie Nutzung publizierten Wissens durch die Akteure in Bildung und Wissenschaft ist Bedingung für die Gemeinfreiheit von Wissen und Information für alle*".

384 *Kuhlen,* Die Transformation der Informationsmärkte in Richtung Nutzungsfreiheit, S. 60 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

III. Die Auswirkungen der Wissenschaftsfreiheit auf die Schrankenauslegung

Basierend auf der Erkenntnis, dass sich die Interessenlage im Wissenschaftsurheberrecht vom sonstigen Urheberrecht unterscheidet, wird im folgenden Abschnitt erläutert, in welchem Umfang sich Wertungen zugunsten der Freiheit von Wissenschaft und Forschung im GG sowie der GRCh auf die Auslegung der Schrankenbestimmungen auswirken. Dafür ist zunächst das Konkurrenzverhältnis zwischen GG und GRCh aufzulösen.

Art. 51 Abs. 1 S. 1 GRCh macht die Beantwortung der Frage danach, ob das deutsche Grundgesetz oder die GRCh anzuwenden ist, davon abhängig, ob das nationale Recht „*vollständig vom Unionsrecht bestimmt*" ist,[385] andernfalls bleiben die nationalen Grundrechte Prüfungsmaßstab. Wann das nationale Recht allerdings vollständig unionsrechtlich bestimmt ist, ist in seinen Details zwischen dem EuGH und dem BVerfG strittig,[386] wobei beide Gerichte jeweils den Vorrang ihrer eigenen Rechtsordnung präferieren.[387] Auch in diesen festen Prinzipien sind jedoch Aufweichun-

---

385 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 936 Rn. 32.

386 Streinz/Michl, EUV/AEUV, Art. 4 EUV Rn. 35; u. a. EuGH, Urt. v. 22. 10. 1998 - Rs. C-10–97 bis C-22–97 – *Erstattung gemeinschaftsrechtswidrig erhobene Abgaben* = NJW 1999, S. 200 Rn. 21. In Bezug auf die Darstellung der Problematik bis zum Jahr 2010 sei auf *Nusser*, Die Bindung der Mitgliedstaaten an die Unionsgrundrechte und seine ausführliche Aufarbeitung der EuGH-Rechtsprechung verwiesen.

387 Der EuGH wendet die GRCh an, wenn der Geltungsbereich des Unionsrechts eröffnet ist, so die st. Rspr.: „*Aus der ständigen Rechtsprechung des* Gerichtshofs *ergibt sich im Wesentlichen, dass die in der Unionsrechtsordnung garantierten Grundrechte in allen unionsrechtlich geregelten Fallgestaltungen, aber nicht außerhalb derselben Anwendung finden. Insoweit hat der Gerichtshof bereits festgestellt, dass er eine nationale Rechtsvorschrift nicht im Hinblick auf die Charta beurteilen kann, wenn sie nicht in den Geltungsbereich des Unionsrechts fällt. Sobald dagegen eine solche Vorschrift in den Geltungsbereich des Unionsrechts fällt, hat der im Rahmen eines Vorabentscheidungsersuchens angerufene* Gerichtshof *dem vorlegenden Gericht alle Auslegungshinweise zu geben, die es benötigt, um die Vereinbarkeit dieser Regelung mit den Grundrechten beurteilen zu können, deren Wahrung er sichert*", EuGH, ECLI:EU:C:2013:105 – *Åkerberg Fransson* = NJW 2013, S. 1415, 1415 f. Rn. 19; *A. Schwerdtfeger* in Meyer/Hölscheidt, EU GRCh, Art. 51 Rn. 40; *J. P. Terhechte* in von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht, Art. 51 GRC Rn. 9. Das BVerfG hingegen wendet die Grundrechte des deutschen GGs dann an, wenn der nationale Gesetzgeber Spielräume lässt, vgl. nur BVerfG, Urt. v. 2.3.2010 - 1 BvR 256/08, 1 BvR 263/08, 1 BvR 586/08 – *Vorratsdatenspeicherung*

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

gen zu registrieren: Auf der einen Seite beschloss das BVerfG in „*Recht-auf-Vergessen-II*“ unter Anerkennung einer Letztentscheidungskompetenz des EuGHs bei Auslegungsfragen, dass es auch selbst die Unionsgrundrechte prüfen kann.[388] Auf der anderen Seite entschied der EuGH explizit zum Urheberrecht, dass erstens die Schrankenbestimmungen der InfoSoc-RL (Art. 5 Abs. 2 und 3) nicht vollharmonisierend sind und insofern zweitens nationale Grundrechte angewandt werden können, sofern die nationalen Grundrechte das Unionsniveau nicht unterschreiten.[389] Deswegen gilt, dass das Wissenschaftsurheberrecht am GG gemessen wird, wenn es Spielräume gibt. Ist dies nicht der Fall, ist die GRCh Maßstab. In Bezug auf die Frage, wann eine Norm vollständig unionsrechtlich bestimmt ist, kommt es nach EuGH-Rechtsprechung im Einzelfall auf den Wortlaut der jeweiligen Bestimmung an.[390] Nicht schädlich ist ausweislich dieser EuGH-Rechtsprechung, dass der Umsetzungsspielraum der nationalen Gesetzgeber begrenzt ist, insbesondere durch den Wortlaut oder den Dreistufentest.[391] Verbleibt ein hinreichender Umsetzungsspielraum für die Mitgliedsstaaten, sind bei der Auslegung der Schranken nationale Grundrechte anzuwenden.[392]

= MMR 2010, S. 356, 358 Rn. 182; *A. Schwerdtfeger* in Meyer/Hölscheidt, EU GRCh, Art. 51 Rn. 43.

388 BVerfG, Beschl. v. 6.11.2019 – 1 BvR 276/17 – *Recht auf Vergessen II* = NJW 2020, S. 314, 318 Rn. 50 ff.; *J. M. Hoffmann*, NVwZ 2020, S. 33, 34; *Kühling*, NJW 2020, S. 275, 279 versteht das als eine „*Neuausrichtung*“ der Rolle des BVerfG in Zusammenspiel mit dem EuGH.

389 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 936 Rn. 28, 30 f.; EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, S. 940, 941 Rn. 19, 20.

390 Z. B. „*soweit es der Informationszweck rechtfertigt*“ und „*sofern die Nutzung den anständigen Gepflogenheiten entspricht und in ihrem Umfang durch den besonderen Zweck gerechtfertigt ist*“, EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, S. 940, 941 Rn. 25; EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 937 Rn. 40; *T. Dreier*, GRUR 2019, S. 1003, 1004.

391 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 937 Rn. 45 ff.

392 Das setzt schließlich auch der BGH um, BGH, Urt. v. 30.4.2020 – I ZR 139/15 – *Afghanistan-Papiere II* = GRUR 2020, S. 853, 857 Rn. 50; BGH, Urt. v. 30.4.2020 – I ZR 228/15 – *Reformistischer Aufbruch II* = GRUR 2020, S. 859, 865 Rn. 61. Dass das Schutzniveau der GRCh nicht unterschritten wurde, nahm der BGH dabei mit dem Verweis an, dass keine gegenteiligen Anhaltspunkte ersichtlich seien.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Jedenfalls sind die Verwertungsrechte der InfoSoc-RL, d. h. das Vervielfältigungsrecht und das Recht der öffentlichen Wiedergabe inklusive des Rechts der öffentlichen Zugänglichmachung, vollharmonisiert.[393] In Bezug auf die Schrankenbestimmungen haben diese Grundsätze zur Folge, dass insbesondere zwischen den verschiedenen Richtlinien, auf denen die nationalen Normen basieren, unterschieden werden muss. In Wissenschaftskontexten ermöglicht Art. 5 Abs. 3 lit. a InfoSoc-RL Ausnahmen oder Beschränkungen von den Verwertungsrechten des Urhebers gem. Art. 2, 3 InfoSoc-RL *„für Zwecke der wissenschaftlichen Forschung, sofern – außer in Fällen, in denen sich dies als unmöglich erweist – die Quelle, einschließlich des Namens des Urhebers, wann immer dies möglich ist, angegeben wird und soweit dies zur Verfolgung nicht kommerzieller Zwecke gerechtfertigt ist"*. Ein Umsetzungsspielraum der Mitgliedsstaaten ergibt sich daher insbesondere in der Beurteilung der Rechtfertigung einer Ausnahme oder Beschränkung durch nicht-kommerzielle Zwecke und der daraus resultierenden Frage, in welchem Umfang sie vorgesehen werden dürfen. Insofern ist Art. 5 Abs. 3 lit. a InfoSoc-RL nicht als Maßnahme anzusehen, die mitgliedsstaatliches Recht vollständig harmonisiert, weswegen die nationalen Grundrechte Maßstab sind.

Die Ausnahmen und Beschränkungen der DSM-RL (speziell in Bezug auf das TDM sind das Art. 3 und Art. 4, für die Lehre gilt Art. 5 und für Kulturerbe-Einrichtungen Art. 6) sind demgegenüber verpflichtend ausgestaltet. Dieser verpflichtende Charakter stellt einen entscheidenden Unterschied zu den Ausnahmen und Beschränkungen der InfoSoc-RL dar. Auf Grundlage der Ausnahmen und Beschränkungen InfoSoc-RL dürfen die Mitgliedsstaaten allerdings über die Vorgaben hinausgehen, Art. 24, 25 DSM-RL. Diese Vorgaben der InfoSoc-RL sind – auch ausweislich der EuGH-Rechtsprechung[394] – wiederum nicht vollharmonisierend. Aus diesem Grund muss das auch für die Ausnahmen und Beschränkungen, die auf der DSM-RL basieren, gelten.[395] Insbesondere in den Bereichen, in denen der nationale Gesetzgeber sich in einer überschießenden Umsetzung der DSM-RL auf die InfoSoc-RL stützt, müssen also nationale Grundrechte angewandt werden.

---

393 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 937 Rn. 37 f.

394 EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 936 Rn. 28, 30 f.; EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, S. 940, 941 Rn. 19, 20.

395 *Stieper*, GRUR 2020, S. 1, 3.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Diskussionsbedürftig bleibt, ob auch Umsetzungsrecht, das sich nahezu vollständig oder vollständig an die Vorgaben der Richtlinie hält, statt Umsetzungsspielräume zu nutzen, an den nationalen Grundrechten zu messen ist. Dafür spricht, dass der nationale Gesetzgeber auch in diesen Bereichen die Abweichungsmöglichkeiten hatte. Nach Art. 51 Abs. 1 S. 1 GRCh muss ein Akt eine „*Durchführung des Rechts der Union*" sein. Nach der EuGH-Rechtsprechung ist dabei auf den Wortlaut der unionsrechtlichen (Ermächtigungs-)Norm bzw. den sich daraus resultierenden Spielraum abzustellen. Das spricht dafür, nicht die tatsächliche Abweichung im Umsetzungsrecht, sondern die theoretische Abweichungsmöglichkeit heranzuziehen.

Diese Erwägungen haben zur Folge, dass in Wissenschaftskontexten die Normen des deutschen GG Maßstab der urheberrechtlichen Schrankenbestimmungen sind. Wegen der Berücksichtigung der wissenschaftlichen Eigengesetzlichkeiten in den sachlichen Schutzbereich der Wissenschaftsfreiheit[396] sind das Selbstverständnis von Wissenschaft und die Anforderungen an wissenschaftliche Arbeitsweisen und Standards einzubeziehen.[397] Als Abwehrrecht bindet die Freiheit von Wissenschaft gem. Art. 1 Abs. 3 GG in erster Linie die staatlichen Organe, insofern auch die Gesetzgebung und die Rechtsprechung.[398] Die Wissenschaftsfreiheit enthält daneben auch eine „*objektive, das Verständnis von Wissenschaft, Forschung und Lehre zum Staat regelnde, wertentscheidende Grundsatznorm*".[399] Dies wird zwar i. d. R. durch die Finanzierung, Organisation sowie die

396 S. dazu bereits o. S. 74 ff.

397 *Pernice* in Dreier, GG, 2. Aufl. 2004, Art. 5 III (Wissenschaft) Rn. 25; *A. L. Paulus* in v. Mangoldt/Klein/Stark, GG, Art. 5 Abs. 3 Rn. 473; anschaulich erläutert das auch *Morlok,* Selbstverständnis als Rechtskriterium, S. 97: „*Der Bezug der Wissenschaftlichkeit bei der Bestimmmung der wissenschaftlichen Bemühungen oder Kommunikationen ist Ausdruck der kommunikativen Geschlossenheit des gesellschaftlichen Teilsystems Wissenschaft und nimmt folgerichtig deren Selbstqualifizierung in Bezug*", zu beachten sind dabei insbesondere die Prozesse, die das wissenschaftliche Arbeiten „*anleiten und inhaltlich prägen*", *Häberle*, AöR 1985, S. 329, 356 f.; vgl. dazu an späterer Stelle S. 268 ff.

398 *Jarass* in Jarass/Pieroth, GG, Art. 5 Rn. 134, Art. 1 Rn. 32, 34.; BVerfG, Beschl. v. 9. 11. 2004 - 1 BvR 684/98 = NJW 2005, S. 1413, 1414.

399 So die st. Rspr. des BVerfG, BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, S. 1176, 1177; BVerfG, Beschl. v. 26. 10. 2004 - 1 BvR 911/0 = NVwZ 2005, S. 315, 316; BVerfG, Beschl. v. 03.03.1993 - 1 BvR 757, 1551/88 = NVwZ 1993, S. 663, 664; BVerfG, Beschl. v. 31.05.1995 - 1 BvR 1379, 2 BvR 1413/94 = NVwZ 1996, S. 709, 710; BVerfG, Beschl. v. 26. 10. 2004 - 1 BvR 911/0 = NVwZ 2005, S. 315, 316; *Jarass* in Jarass/Pieroth, GG, Art. 5 Rn. 133.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Ausgestaltung des Wissenschaftsbetriebs verstanden.[400] In Konflikt mit der Gewährleistung des Eigentums, Art. 14 Abs. 1 GG, ergibt sich allerdings auch der staatliche Auftrag, die Grundrechte durch eine praktische Konkordanz in einen möglichst schonenden Ausgleich zu bringen.[401] Andernfalls sind je nachdem die Freiheitsrechte der betroffenen Nutzer verletzt.[402]

Bei der Vornahme der praktischen Konkordanz bzw. der Abwägung der sich gegenüberstehenden Grundrechte kommt dem Gesetzgeber ein nicht unerheblicher Gestaltungsspielraum zu.[403] Je stärker eine wissenschaftliche Tätigkeit den Kernbereich der Wissenschaftsfreiheit betrifft, desto eher muss eine Abwägung zugunsten der Wissenschaftsfreiheit ausfallen. Ein angemessener Ausgleich für die Träger des gem. Art. 14 GG geschützten Eigentums kann zum einen darüber erfolgen, dass die Inhalts- und Schrankenbestimmungen nur in geringerem Umfang zugelassen werden und zum anderen dadurch, dass sie vergütungspflichtig ausgestaltet werden.

Auch die Judikative hat die Wertungen der Grundrechte zu berücksichtigen, wenn sie die Schrankenbestimmungen zugunsten der Wissenschaft auslegt: Dem kann auch nicht entgegengehalten werden, dass diese Berücksichtigung schon aufgrund einer notwendigerweise eng zu erfolgenden Schrankenauslegung nicht möglich sei.[404] Eine weitere Eingrenzung ist spätestens seit der Rechtsprechung des EuGHs in den Sachen „*Funke Medien/Bundesrepublik Deutschland*“, „*Spiegel Online/Volker Beck*“ sowie „*Pelham/Hütter*“ gezogen werden: Aus den genannten Entscheidungen ergibt sich, dass eine grundrechtskonforme Auslegung jedenfalls nur innerhalb vorhandener Schranken in Betracht kommt, weil die Schranken der

---

400 BVerfG, Urt. v. 29. 5. 1973 - 1 BvR 424/71 u. 325/72 = NJW 1973, 1176, S. 1177; *von der Decken* in Schmidt-Bleibtreu/Hofmann/Henneke, GG, Art. 5 Rn. 47; *Paul Baumann/Krahn/Lauber-Rönsberg,* Forschungsdatenmanagement und Recht: Datenschutz-, Urheber- und Vertragsrecht, S. 19.

401 *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 46; OLG München, Urt. v. 09.03.1995 - 29 U 3903/94 – *"Telefon-Sex-Foto"* = NJW-RR 1996, S. 539; *P. Raue* in FS Nordemann, S. 327, 337, bei der praktischen Konkordanz gilt, dass die Grundrechte gegeneinander so abgewogen werden müssen, dass sie beide die größte Wirksamkeit entfalten, *K. Hesse,* Grundzüge des Verfassungsrechts der Bundesrepublik Deutschland Rn. 72, 317 f.

402 *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 46.

403 *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 47; BVerfG, Beschl. v. 11.10.1988 - 1 BvR 743/86 = NJW 1992, S. 1307, 1308.

404 S. dazu S. 79 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

InfoSoc-RL abschließend geregelt sind.[405] Es ist jedenfalls keine „*freischwebende Grundrechtsabwägung*" möglich.[406] Bei der grundrechtskonformen Auslegung innerhalb der Schranken sind gleichwohl die Zielsetzungen der jeweiligen Bestimmungen bzw. der jeweiligen Richtlinie zu beachten.[407]

Noch nicht beantwortet ist jedoch die Frage danach, wann bestehende Schranken grundrechtskonform ausgelegt werden müssen, grundrechtliche Wertungen also Anklang in der Auslegung finden. *Stieper* und *Martens* beschränken die grundrechtskonforme Schrankenauslegung auf „*offensichtliche Verfehlungen*" des Grundrechtsschutzes.[408] *Stieper* begründet das damit, dass die Schrankentatbestände bereits eine Ausgestaltung der grundrechtlichen Interessenabwägung darstellen und insofern nicht jeder Schrankenprüfung eine weitere Interessenabwägung nachgeschaltet werden dürfe.[409] Diesen Forderungen ist dahingehend zuzustimmen, dass der

---

405 EuGH, ECLI:EU:C:2019:624 – *Pelham/Hütter* = GRUR 2019, S. 929, 932 Rn. 58 f.; EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, 940, S. 943 Rn. 42 ff.; EuGH, ECLI:EU:C:2019:623 – *Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere]* = GRUR 2019, S. 934, 939 Rn. 57 ff. So bereits vor den genannten Entscheidungen *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 64, 69 f.; *Stieper* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, Vor § 44a ff. Rn. 39; BGH, Urt. v. 20. 3. 2003 - I ZR 117/00 – *Gies-Adler* = GRUR 2003, S. 956, 657; trotz inhaltlicher Zurückhaltung in der Schrankenauslegung auch BGH, Urt. v. 24.1.2002 - I ZR 102/99 – *Verhüllter Reichstag* = GRUR 2002, 605, S. 606; BVerfG, Beschl. v. 29. 6. 2000 - 1 BvR 825/98 – *Grenzen der Zitierfreiheit (Germania 3)* = GRUR 2001, S. 149, 151. Anderer Auffassung ist *Findeisen,* Die Auslegung urheberrechtlicher Schrankenbestimmungen, S. 132.

406 *T. Dreier*, GRUR 2019, S. 1003 ff., 1005; *Leistner*, ZUM 2019, S. 720 ff., 725; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 64.

407 U. a. EuGH, ECLI:EU:C:2019:625 – *Spiegel Online/Volker Beck [Reformistischer Aufbruch]* = GRUR 2019, S. 940, 942 Rn. 35; EuGH, ECLI:EU:C:2011:798 – *Painer/Standard* = GRUR 2012, S. 166, 169 Rn. 107.

408 *Stieper*, GRUR 2017, S. 1209, 1212; *S. A. E. Martens,* Methodenlehre des Unionsrechts, S. 529.

409 *Stieper*, ZUM 2020, S. 753, 760; *Stieper,* Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 64 f. Kritisch diesbezüglich *Findeisen,* Die Auslegung urheberrechtlicher Schrankenbestimmungen, S. 118, 124: „*Nun wurde im Bereich des Urheberrechts lange Zeit nahezu einmütig davon ausgegangen, die Schrankenbestimmungen stellten bereits das Ergebnis eines vom Gesetzgeber vorgenommenen Abwägungsprozesses dar, in welchem dieser die verfassungsrechtlich geschützten Interessen des Urhebers mit den widerstreitenden Interessen des Urhebers mit den wderstreitenden Allgemeinwohlbelangen bereits zu einem angemessenen, grundsätzlich abschließenden ausgleich gebracht habe*".

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Gesetzgeber bei der Normierung der Schrankenbestimmungen seinen Gestaltungsspielraum bereits dahingehend nutzte, die Grundrechte gegeneinander in Ausgleich zu bringen. Diese Wertentscheidung darf nicht durch eine grundrechtskonforme Auslegung untergraben werden. Bei der Frage, ob Normen des Wissenschaftsurheberrechts grundrechtskonform auszulegen sind, kann dabei in Abstufungen differenziert werden, d. h. Schranken sind, wenn sie Vorgänge erfassen, die eine wissenschaftliche Arbeit erst als wissenschaftlich definieren, eher einer grundrechtskonformen Auslegung zu unterwerfen als Prozesse, die zwar wissenschaftlich erwünscht sind, aber nicht notwendig für die Anerkennung der Forschung sind.

D. Zwischenfazit

Das TDM kann als Methode der Wissenschaft fungieren und unterliegt deswegen den besonderen Wertungen des Wissenschaftsurheberrechts. Der Systematik des Urheberrechts ist eine Privilegierung der Wissenschaft immanent – freies Wissen soll die Monopolisierung wissenschaftlicher Information verhindern[410] und den Fortschritt begünstigen. Dabei gilt: Je wissenschaftlicher ein wissenschaftliches Erzeugnis ausfällt, desto niedriger ist das Schutzniveau angesiedelt. Das läuft den Interessen wissenschaftlicher Urheber auch im Grundsatz nicht entgegen, denn diese sind wesentlich anders gelagert als diejenigen vieler anderer Urheber. Insgesamt bildet sich dabei das Bild heraus, dass wissenschaftliche Urheber ihre Geisteserzeugnisse in viel größerem Umfang der Allgemeinheit zur Verfügung stellen, als es in anderen Bereichen urheberrechtlichen Werkschaffens der Fall ist. Das wirft die Frage nach den Schrankenbestimmungen auf, die weder Ausnahmebestimmungen, noch grundsätzlich eng auszulegen sind, sondern berechtigte Interessen privilegieren und auf eine gesetzgeberische Entscheidung zurückgehen, bei der das geistige Eigentum mit der Freiheit von Wissenschaft und Forschung abzuwägen ist. Um wissenschaftliche Werke schaffen zu können bzw. neue Erkenntnisse zu generieren, ist die Wissenschaft allerdings im Gegenzug auch darauf angewiesen, auf vorbestehendem Wissen aufbauen zu können. Hierfür sind Schrankenbestimmungen erforderlich, die in dieser Arbeit in Bezug auf das TDM eine

410 *Kuschel*, OdW 2020, S. 43, 43; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 41, 93; *Lutz,* Zugang zu wissenschaftlichen Informationen in der digitalen Welt, S. 45 f.; *v. Moltke,* Das Urheberrecht an den Werken der Wissenschaft, S. 6, 74 f.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

besondere Beachtung finden werden und bei deren Auslegung innerhalb gewisser Grenzen auch Grundrechtswertungen zu berücksichtigen sind.

§ 5 Fortschritt durch Restriktion oder Restriktion des Fortschritts? – Die urheberrechtliche Steuerung von Innovation

Das Urheberrecht nimmt im Grundsatz keine Unterscheidung zwischen traditionellen und modernen, analogen und digitalen Sachverhalten vor. In den vergangenen 20 Jahren sind allerdings punktuelle Reformen erfolgt, die technische Vorgänge ermöglichen und damit technologischen Fortschritt begünstigen sollten. Dazu zählt auch das TDM: Mithilfe spezifischer Erlaubnisnormen auf deutscher und europäischer Ebene sollte das TDM als Schlüsseltechnologie in wissens- und sprachbasierten Systemen sowie in der Erforschung und Entwicklung von Künstlicher Intelligenz unterstützt und wissenschaftlicher wie technischer Fortschritt gefördert werden.[411] Infrage steht deswegen, wie die betreffenden Akteure am besten dazu motiviert werden können, TDM vorzunehmen und die Technologie fortzuentwickeln. Deswegen adressiert der folgende Abschnitt das viel diskutierte Problem, inwiefern das Urheberrecht als Anreiz für Innovation fungieren kann. Lange Zeit wurde angenommen, dass das Urheberrecht jedenfalls für traditionelles Werkschaffen Anreize setzt, indem es ein hohes Schutzniveau herstellt. Die daraus resultierenden Annahmen wurden auch auf Innovationen übertragen. Unter „*Innovation*“ wird weitläufig eine Neuerung oder Veränderung bzw. Reform aus verschiedenen Bereichen, insbesondere eine neue Technik, verstanden.[412] Das wirft die besondere Problematik auf, dass im Bereich von Innovation heterogene – sowohl ökonomische als auch wissenschaftliche Interessen – betroffen sind. Das wird im Folgenden kritisch beleuchtet. Außerdem werden weitere Ansätze, wie das Urheberrecht technologiefreundlich verstanden und fortgebildet werden kann, diskutiert.

---

411 Zum deutschen Recht die Bundesregierung, Strategie Künstliche Intelligenz der Bundesregierung (2018), S. 8; Unterrichtung durch die Bundesregierung zur Strategie Künstliche Intelligenz - Fortschreibung 2020, BT-Drs. 19/25095, S. 8 ff.; zum europäischen Recht Europäische Kommission, Künstliche Intelligenz für Europa - COM(2018) 237 final, S. 3 ff.; Europäische Kommission, Weißbuch zur Künstlichen Intelligenz - COM(2020) 65 final, S. 1.

412 So der Duden, Definition der Innovation, https://www.duden.de/rechtschreibung/Innovation.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

A. Das urheberrechtliche Anreizparadigma

Dem traditionellen Urheberrechtsverständnis liegt der Gedanke zugrunde, dass mittels eines möglichst hohen Schutzniveaus kreatives Werkschaffen motiviert werden kann.[413] Auf dieser Annahme basiert nicht nur das traditionelle Urheberrecht, sondern auch das Urheberrecht als Recht der Digitaltechnik sowohl in Bezug auf das Urheberrecht als auch in Bezug auf Leistungsschutzrechte.[414] Der InfoSoc-RL liegt etwa die Erwägung zugrunde, dass eine „*erhöhte Rechtssicherheit und […] die Wahrung eines hohen Schutzniveaus im Bereich des geistigen Eigentums substantielle Investitionen in Kreativität und Innovation einschließlich der Netzinfrastruktur fördern und somit zu Wachstum und erhöhter Wettbewerbsfähigkeit der europäischen Industrie beitragen*“ könne.[415] Auch die DSM-RL, die im Vergleich zur InfoSoc-RL Schranken ein größeres Gewicht verleiht, geht davon aus, dass ein hohes Schutzniveau gewahrt werden müsse.[416]

Aus ökonomischer Perspektive existieren für die Gewährung eines hohen Schutzniveaus zwei Begründungsansätze, wobei der erste eine Unterproduktion von Immaterialgütern aufgrund eines zu niedrigen Schutzniveaus (sog. *tragedy of commons*) verhindern und der zweite eine Verkehrsfähigkeit immaterieller Güter herstellen möchte, um einem Marktversagen

---

413 *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261, 261 f.; *Kerber* in: Grünberger/Leible (Hrsg.), Die Kollision von Urheberrecht und Nutzerverhalten, S. 31, 32 f. Das betrifft einerseits die urheberpersönlichkeitsrechtliche Komponente des Urheberrechts und andererseits den Umfang an erteilten Verwertungsrechten, d. h. eine möglichst große wirtschaftliche Beteiligung am eigenen Geisteswerk. Aus rein ökonomischen Gesichtspunkten wären nur die Verwertungsrechte zu berücksichtigen, vgl. dazu *Hoffmann-Riem* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 15, 26, der innovative Tätigkeiten von wirtschaftlichen Vorteilen, d. h. verwertungsrechtlichen Rechtsfolgen des anerkannten Immaterialgüterschutzes abhängig macht.

414 Gerade die gesteigerten technischen Neuerungen in Bezug auf Verwertung von Schutzgegenständen verhindere die Amortisierung von Innovationen und Investitionen in die Schaffung urheberrechtlich oder leistungsschutzrechtlich geschützter Gegenstände, was ein Schutzbedürfnis begründe, *Schulze* in FS Schricker, S. 523, 534; *Kur*, GRUR 1990, S. 1, 7 mit Verweis auf die zunehmende Ökonomisierung des Urheberrechts durch die Leistungsschutzrechte und die wachsende Anreizfunktion mittels Investitionssicherung; ähnlich Hilty, zusammengefasst von *Bassewitz/B. Kramer/Prinz zu Waldeck und Pyrmont*, GRUR Int. 2004, S. 607, 608.

415 Richtlinie (EU) 2001/29/EGI, InfoSoc-RL, Erw. 4.

416 DSM-RL Erw. 9.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

zu begegnen.[417] Schranken haben gerade im letztgenannten Begründungsansatz die Funktion, Marktversagen zu korrigieren.[418]

Die Annahme, die Gewährung ausschließlicher Verwertungsbefugnisse könne Anreize setzen, ist auch im Technikrecht von langer Tradition geprägt: Schon die Drucker- und Gewerbeprivilegien, die eher mit dem Patent- als mit dem Urheberrecht vergleichbar sind,[419] dienten als Investitionsschutz der innovationspolitischen Anreizsetzung.[420] Auch dem modernen Patentrecht wird dieses Ziel zugesprochen.[421] Empirisch ist jedoch nicht geklärt, ob ökonomische Anreize Kreativität fördern können.[422]

---

417 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 14; ausführlich *Leistner*, ZGE 2009, S. 403, 406 ff.; *Leistner/Hansen*, GRUR 2008, S. 479, 483 f.; *Bisges*, ZUM 2014, S. 930, 934 f.

418 *Leistner*, ZGE 2009, S. 403, 413.

419 *Ann*, GRUR Int. 2004, S. 597, 600; *Kohler,* Urheberrecht an Schriftwerken und Verlagsrecht § 5: *„gewerbliches Monopol"*; vgl. auch *Jänich,* Geistiges Eigentum - eine Komplementärerscheinung zum Sacheigentum?, S. 21 f.

420 *Kuschel* in: Eifert/Hoffmann-Riem (Hrsg.), Digitale Disruption und Recht, S. 93, 94 f.; *Ann*, GRUR Int. 2004, S. 597, 600; *Seifert*, NJW 1992, S. 1270, 1273 f.; ausführlich *Jänich,* Geistiges Eigentum - eine Komplementärerscheinung zum Sacheigentum?, S. 20 ff.; Ann in *Ann,* Patentrecht § 4 Rn. 3; *Götting,* Gewerblicher Rechtsschutz § 2 Rn. 13; *Osterrieth,* Patentrecht Rn. 78; *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 11; grundlegend zum Privilegienwesen *Heymann* in: Preußische Akademie der Wissenschaften (Hrsg.), Sitzungsberichte der preußischen Akademie der Wissenschaften, Gesamtsitzung am 24. März, Mitteilung aus der Sitzung der phil.-hist. Klasse vom 2. Dezember 1926, S. 49, 65 ff.

421 *Kuschel* in: Eifert/Hoffmann-Riem (Hrsg.), Digitale Disruption und Recht, S. 93, 94; diese Annahme wird heute auch teils differenzierter betrachtet, vgl. *Ann* in Ann, Patentrecht § 3 Rn. 39 zufolge kann weder das Ausbleiben technischischer Neuerungen ohne einen Patentschutz auf dessen Fehlen zurückgeführt werden, noch kann begründet werden, der Patentschutz fungiere überhaupt nie als Anreiz; *Kirchner* in: Ott/Schäfer (Hrsg.), Ökonomische Analyse der rechtlichen Organisation von Innovationen, S. 157, 170 ff. auf Grundlage einer Kosten-Nutzen-Abwägung, derzufolge es jedenfalls alternativ zu einem Patentschutz ein anderes Ausschließlichkeitsrecht oder auch Geheimhaltungsrecht bedürfte, um Produktinnovationen überhaupt zu ermöglichen; *Harabi* in: Harabi (Hrsg.), Kreativität - Wirtschaft - Recht, S. 61, 89 ff.; auf Grundlage eines Berechnungsmodells hegt *Engel* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 43 ff. Zweifel; differenzierter *Bechtold*, GRUR Int. 2008, S. 484, 485 f. demzufolge jedenfalls nicht jede Erhöhung von Schutzrechten zu einer *„erhöhten kreativen oder erfinderischen Tätigkeit"* führt.

422 *Reich,* Die ökonomische Analyse des Urheberrechts in der Informationsgesellschaft, S. 33 f.; *Leistner*, ZGE 2009, S. 403, 417 f. weist auch darauf hin, dass sich z. B. das Patentrecht viel eher eignet, um Anreizthesen zu prüfen, so sind schließlich Patentanmeldungen einfach zu messen. Für das Urheberrecht

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Bedenken an einer strikten Ableitung von Verhaltensweisen aus ökonomischen Vorteilen kommen erst recht wegen der ideellen, urheberpersönlichkeitsrechtsbasierten, Elementen des Urheberrechts auf.[423] Aus diesem Grund wird inzwischen angenommen, dass hinsichtlich traditioneller Werkschaffender statt eines möglichst hohen Schutzniveaus eine Kompensations- bzw. Ausgleichsfunktion genügt und finanzielle Nachteile von Werkschaffender das Werkschaffen nachteilig beeinflussen.[424] Dasselbe gilt aber für ein zu hoch angesiedeltes Schutzniveau, bei dem das Werk faktisch nicht mehr von Dritten nutzbar ist.[425]

Die ideellen Bestandteile des Urheberrechts wirken sich auch im Bereich wissenschaftlichen Werkschaffens aus, denn aus wissenschaftssoziologischer Perspektive sind intrinsische und ideelle Motive ein wichtiger Beweggrund für Geistesschaffen, insbesondere altruistisches wissenschaftliches Erkenntnisinteresse oder künstlerisch-gestaltende Betätigung.[426] Ent-

---

schlägt er insbesondere spieltheoretische Untersuchungen vor. Allgemein zum geistigen Eigentum untersucht das *Engel*, J. Competition Law Econ. 2011, S. 277 ff.

423 *Reich*, Die ökonomische Analyse des Urheberrechts in der Informationsgesellschaft, S. 38 f.; differenzierend *Ohly* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 279, 288, der sowohl ideelle als auch monetäre Interessen einbezieht; ebenso *Geiger*, IIC 2006, S. 371, 380, der kritisiert, dass das Werkschaffen nicht allein auf einen ökonomischen Prozess reduziert werden dürfe; einen psychologischen Ansatz zur Untersuchung des Einflusses von Kreativität und Anreizen verfolgt *Peukert* in: Rehbinder (Hrsg.), Die psychologische Dimension des Urheberrechts, S. 113, 138 ff., der – intrinsische und extrinsische Motive differenzierend – intrinsische Motive als Grundvoraussetzung betrachtet, extrinsische Motive aber als kreativitätsfördernd einordnet. Gleichwohl ist dieser Bereich der Kreativitätsforschung aus Erkenntnisperspektive stark begrenzt. Zu der Problematik, Mechanismen, die in Bezug auf ein Schutzrecht bestehen, auf ein anderes zu übertragen *Peukert* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 299, 301 f.

424 Für eine Amortisierung der Fixkosten *Bechtold*, GRUR Int. 2008, S. 484, 485; ebenso *Geiger*, IIC 2006, S. 371, 380; aus US-amerikanischer Perspektive *Frischmann/Lemley*, Colum. L. Rev. 2007, S. 257, 276.

425 *T. Dreier*, GRUR Int. 2015, S. 648, 650.

426 *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261, 265 f. in Bezug auf nicht-kommerzielle Gedichte oder Musikstücke, aber auch auf Werke, die im Rahmen von Arbeitseinkommen oder Stipendien geschaffen werden, d. h. insbesondere auch Werke der Wissenschaft; *Peukert*, JIPITEC 2013, S. 142, 145 Rn. 20; die Relevanz ökonomischer Aspekte in weiten Bereichen nichtwissenschaftlicher Publikationen betont *Brintzinger*, Leviathan 2010, S. 331, 345; auch *Schack* in: Depenheuer/Peifer (Hrsg.), Geistiges Eigentum: Schutzrecht oder Ausbeutungstitel?, S. 123, 134 betont die Bedeutung intrinsischer Motive, obwohl er auch einräumt, das im Bereich mancher

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

sprechend der dargestellten Interessen wissenschaftlicher Urheber – jedenfalls derjenigen, die ihren Lebensunterhalt als Arbeitnehmerurheber bestreiten – steigern wissenschaftliche Publikationen die Reputation, sie stellen aber aus finanzieller Hinsicht allenfalls eine Nebeneinnahmequelle dar[427] und stehen in vielen Fällen gerade nicht in unmittelbarem Zusammenhang mit weitreichenden Verwertungsbefugnissen. Es kommt also weniger auf den Umfang der Zuordnung der Verwertungsrechte als auf die persönlichkeitsrechtlichen Bestandteile des Urheberrechts an, insbesondere die Anerkennung als Urheber.[428] Plausibel erscheint die Anreizthese insofern höchstens in Bezug auf Urheber, die keine Arbeitnehmerurheber sind, sowie deren Verwerter.[429]

Daneben übersieht die Annahme, Werkschaffen werde durch ein hohes Schutzniveau motiviert, dass das Werkschaffen – gerade wenn es um wissenschaftlichen Fortschritt und Innovation geht – darauf basiert, auf vorhandenem (urheberrechtlichen) Schutzgut aufzubauen, wofür die Nutzung bereits vorhandenen Wissens, das oft in urheberrechtlich geschützter Form publiziert wurde, notwendig ist.[430] Ein Schutzniveau, das zu wenige oder zu eingeschränkte Nutzungserlaubnisse enthält, kann sich demnach sogar nachteilig auf das Geistesschaffen auswirken und – statt Anreize zu setzen – Fortschritt behindern. Das Marktversagen, das ursprünglich durch ein hohes Schutzniveau verhindert werden sollte, wird auf diese Weise gerade herbeigeführt.[431]

Werke, die in ihrer Erstellung besonders teuer sind, Anreize notwendig sind, die finanzielle Erträge dem Inhaber des Rechts zuordnen; schon Werner von Humboldt erkannte die Bedeutung intrinsischer Motive, vgl. *Humboldt*, Schriften zur Politik und zum Bildungswesen, so betont er das „*geistige Leben des Menschen*“, die „*äußere Muße*“, aber auch das „*innere Streben zur Wissenschaft*“; aus wissenschaftssoziologischer Perspektive auch *Merton* in: Weingart (Hrsg.), Wissenschaftliche Entwicklung als sozialer Prozeß, S. 45, 53 ff.

427 S. dazu bereits S. 101 f.

428 S. dazu bereits S. 101 f.

429 *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261, 266 weist darauf hin, dass dies gerade die Personengruppe darstellt, die vom historischen Gesetzgeber bedacht wurde.

430 *Haedicke* in: Leible/Ohly/Zech (Hrsg.), Wissen - Märkte - Geistiges Eigentum, S. 93, 100. Das gilt daneben auch für andere Bereiche des Werkschaffens: Der Inbegriff dieses Prinzips war seit jeher die freie Benutzung, § 24 UrhG a. F., die nunmehr durch § 23 sowie § 51a UrhG abgedeckt wird. Zu bemerken ist in diesem Kontext, dass § 24 noch erforderte, dass das neu Geschaffene ein urheberrechtliches Werk war. § 51a UrhG erfordert dies nun nicht mehr.

431 *Bisges*, ZUM 2014, S. 930, 937.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Inzwischen hat sich diese Annahme, d. h. dass sich ein Übermaß an Schutzniveau negativ auf das Werkschaffen, auswirken kann (sog. *tragedy of anticommons*), es also auf den richtigen Zuschnitt der immaterialgüterrechtlichen Schutzrechte, darunter auch des Urheberrechts ankommt, weitestgehend durchgesetzt.[432] Insgesamt ist auf die besondere Bedeutung von Beschränkungen hinzuweisen, die sich gerade als Anreize auswirken können.[433] Doch auch diese Anknüpfung muss aufgrund der Heterogenität von Interessen und der Komplexität kreativen Handelns Vorsicht begegnen.[434]

B. Keine Übertragbarkeit der Anreizthese auf Innovation

Ob und wie das Urheberrechts Innovationen gerade in digitalen Kontexten steuert, ist, wie bereits erläutert, zum einen in Bezug auf die Wirtschaft, die die Innovationen entwickelt und von ihr profitiert, und zum anderen in Bezug auf die Wissenschaft, die durch ihre Forschung Innovation erst ermöglicht,[435] zu thematisieren. Innovationen sind jedoch im Urheber-

---

432 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 14; *T. Dreier*, GRUR Int. 2015, S. 648, 650; ausführlich *Leistner*, ZGE 2009, S. 403, 406 ff.; *Leistner/Hansen*, GRUR 2008, S. 479, 483 f.; aus ökonomischer Perspektive *Scotchmer*, Journal of Economic Perspectives 1991, S. 29, 29 f.; sehr eindrücklich formuliert das *Schack* in FS Wandtke, S. 9, 10 f.: „*Der folgende Beitrag will nicht die Axt an das Urheberrecht legen, sondern aufzeigen, wie man diesen wild gewachsenen Obstbaum so zurückschneiden kann, dass er weniger Schatten wirft, bessere Früchte trägt und wieder mehr geschätzt und respektiert werden kann. Die gegenwärtige Diskussion beschränkt sich regelmäßig auf die Einführung neuer oder die Erweiterung bestehender Schranken. Doch ist es mit einer solchen Schrankenkosmetik, dem Abschneiden einzelner äußerer Zweige, nicht getan. Um den Baum wieder in eine überzeugende Form zu bringen, müssen vielmehr ganze Äste entfernt werden. Gefragt ist also die Säge und nicht die Heckenschere. Meine These […] lautet: Weniger Urheberrecht ist mehr*“; *Bisges*, ZUM 2014, S. 930, 937 mit Hinweis auf Angebot und Nachfrage; mit Verweis auf die Notwendigkeit der Beschränkung von Immaterialgüterrechten *Bechtold*, GRUR Int. 2008, S. 484, 485 f.; allgemein zum geistigen Eigentum *Engel*, J. Competition Law Econ. 2011, S. 277 ff. mit dem Ergebnis, dass die Erteilung eines geistigen Eigentums manchmal, aber nicht immer einen Anreiz für wirtschaftlich-innovatives Tätigwerden bietet.

433 *T. Dreier*, GRUR Int. 2015, S. 648, 650.

434 *Kerber* in: Grünberger/Leible (Hrsg.), Die Kollision von Urheberrecht und Nutzerverhalten, S. 31, 33, 41 f., 49 sieht die Beweggründe insgesamt als zu komplex an, um sie auf Anreize zurückzuführen.

435 Daneben können diese Personenkreise natürlich auch gemeinsam innovativ arbeiten, d. h. in Public-Private-Partnerschaften.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

recht nicht ausdrücklich adressiert, denn Neuheit ist gerade keine urheberrechtliche Schutzvoraussetzung.[436] Ob die Erzeugnisse urheberrechtlichen Geistesschaffens oder urheberrechtlicher Nutzungshandlungen als Innovation oder innovativ bezeichnet werden können, muss folglich individuell entschieden werden.

Etwas anderes gilt für Datenbanken und Computerprogramme, von denen Letztere zwar Werke sind und nur Erstere als wirtschaftliches Leistungsschutzrecht eingeführt wurde, die aber beide nichts mehr mit traditionellem urheberrechtlichem Werkschaffen gemeinsam haben.[437] Doch kann daraus nicht geschlussfolgert werden, dass das Urheberrecht nicht zur Ermöglichung oder Regulierung herangezogen werden kann.[438] Auch das TRIPS-Abkommen geht davon aus, dass das Urheberrecht als Teil des geistigen Eigentums innovationsrechtlich relevant ist,[439] so besagt es in Art. 7: „*Der Schutz und die Durchsetzung von Rechten des geistigen Eigentums sollen zur Förderung der technischen Innovation sowie zur Weitergabe und Verbreitung von Technologie beitragen, dem beiderseitigen Vorteil der Erzeuger und Nutzer technischen Wissens dienen, in einer dem gesellschaftlichen und wirtschaftlichen Wohl zuträglichen Weise erfolgen und einen Ausgleich zwischen Rechten und Pflichten herstellen*".

Insofern gilt es. zu prüfen, ob ein hohes Schutzniveau oder sonstige Mechanismen des Urheberrechtssystems im Bereich von Innovationen, sofern sie denn urheberrechtlich relevant sind, Anreize setzen kann. Teils wird die Annahme, ein hohes Schutzniveau setze diese Anreize, die schon in Bezug auf traditionelles Werkschaffen defizitär ist, auf Innovation übertragen. Insbesondere mit der InfoSoc-RL wurde bezweckt, Innovation mittels eines hohen Schutzniveaus zu fördern.[440] Zudem wurde ein immaterialgüterrechtlicher Schutz im Kontext von KI gefordert, um Anreize für Innova-

436 *Hoffmann-Riem* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 15, 18; *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261, 263; *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022, § 2 Rn. 17.

437 Ähnlich *Ohly* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 279, 281.

438 *Hoffmann-Riem* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 15, 18.

439 Kritisch *Ohly* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 279, 279 f., demzufolge das diesen Eindruck hinterlässt.

440 S. dazu bereits o. S. 114 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

tionen im KI-Bereich zu setzen, um wiederum Investitionen lohnenswerter zu machen und eine Vermarktung zu ermöglichen.[441]

Gegen das Ableiten ökonomischer Anreize aus einem hohen Schutzniveau sprechen verschiedene Aspekte. Zunächst liegt Innovation eine heterogene Interessenlage zugrunde. Gerade die Wissenschaft agiert vorwiegend nach urheberpersönlichkeitsrechtlichen Motiven, d. h. Ziel ist wissenschaftliche Anerkennung, zumeist aber nicht die wirtschaftliche Verwertung. Wissenschaftliche Betätigung hat insofern nicht das Ziel der Inhaberschaft von Verwertungsrechten.

Auch ökonomische Motive gehen nicht immer in der Zubilligung von Verwertungsrechten auf, wie das auf der Datenbank-RL basierende Datenbankherstellerrecht beweist: Dem Datenbankherstellerrecht, §§ 87a ff. UrhG, liegen als Leistungsschutzrecht keine urheberpersönlichkeitsrechtlichen Elemente zugrunde, die Datenbankhersteller profitieren folglich ausschließlich von den wirtschaftlichen Vorteilen. Bezweckt ist, mittels Exklusivität der Datenbanken Anreize für deren Erschaffung zu setzen.[442] Auch im deutschen Umsetzungsrecht setzt sich das Ziel des Investitionsschutzes fort.[443] Im Gegensatz zu wissenschaftlichen Werken vermag es das Datenbankherstellerrecht außerdem, Information mittelbar in gewissem Umfang zu monopolisieren.[444] Im US-amerikanischen Recht sind Datenbanken hingegen – jedenfalls aus urheberrechtlicher Perspektive – nur dann geschützt, wenn sie Werke darstellen, denn ein mit dem europäischen Datenbankherstellerrecht vergleichbares Schutzrecht existiert hin-

441 Zu KI-Trainingsdaten *Hacker*, GRUR 2020, S. 1025, 1026 f.; *Hacker*, ZGE 2020, S. 239, 258 f. fordert eine Beseitigung bestimmter Risiken, z. B. bzgl. der Qualität von KI-Produkten aufgrund einer fehlenden Qualität von KI-Trainingsdaten, indem Anreize durch die Gewährung eines Schutzrechts gesetzt werden. Er räumt aber ein, dass auch Schranken – gerade Art. 4 DSM-RL – als Anreize fungieren können; zum Schutz von Erzeugnissen von KI-Systemen *Dornis*, GRUR 2019, S. 1252, 1258; *Zech,* Information als Schutzgegenstand, S. 318, der die Sinnhaftigkeit von Anreizen durch die Zubilligung von Schutzrechten, z. B. Leistungsschutzrechten, von dem Aufwand abhängig macht, den die Erstellung mit sich bringt; *Hilty/J. Hoffmann/Scheuerer* in: Hilty/Lee/Liu (Hrsg.), Artificial intelligence and intellectual property, S. 50, 71 f. jedenfalls in Bezug auf manchen KI-Output, nicht aber auf den Schutz der Modelle selbst.

442 *C. König* in: Hennemann/Sattler (Hrsg.), Immaterialgüter und Digitalisierung, S. 89, 97 f.

443 Richtlinie (EU) 96/6/EG, Datenbank-RL Erw. 7 ff.; *v. Lewinski* in Loewenheim UrhR-HdB, § 60 Rn. 40.

444 Vgl. dazu bereits die Erläuterungen auf S. 90 ff.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

gegen nicht.[445] Trotz des amerikanischen Copyright-Systems, in dem wirtschaftlichen Erwägungen eine viel größere Bedeutung zukommt als im Urheberrechtssystem der EU, wurde schließlich davon abgesehen, zugunsten von Datenbanken ein zusätzliches Schutzrecht zu schaffen.[446] Das hatte gleichwohl keine negativen Auswirkungen auf die Wettbewerbsfähigkeit US-amerikanischer Akteure im Datenbankmarkt.[447] Daran zeigt sich erneut, dass der richtige Zuschnitt der Schutzrechte entscheidend ist, denn auch Datenbankhersteller profitieren gerade von den Befugnissen, Vorbestehendes für die Erstellung ihrer Datenbanken zu verwenden, wozu auch Begrenzungen des Urheberrechts dienen können.[448]

C. Mechanismen zur urheberrechtlichen Innovationsförderung

Auch der deutsche Gesetzgeber sowie der europäische Normgeber gewichten die Schrankenbestimmungen immer stärker: Während die InfoSoc-RL Innovation durch *„erhöhte Rechtssicherheit und die Wahrung eines hohen Schutzniveaus im Bereich des geistigen Eigentums“* fördern sollte,[449] erkannte das UrhWissG im Jahr 2017 ausdrücklich erlaubte Nutzungen als innovationsförderlich an.[450] Noch deutlicher wurde Innovation im Zusammen-

445 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, vor § 87a Rn. 11 f.; zu den völkerrechtlichen Hintergründen *Leistner,* Der Rechtsschutz von Datenbanken im deutschen und europäischen Recht, S. 24 f.

446 Doch wurde dort über die Notwendigkeit eines Datenbankherstellerrechts diskutiert, vgl. U. S. Copyright Office, Report on legal protection for databases (August 1997), S. 72, 88 ff. Pläne eines völkerrechtlichen Abkommens in Bezug auf entsprechende Datenbanken sind auch gescheitert, vgl. dazu *v. Lewinski*, GRUR Int. 1997, S. 667, 680.

447 *C. König* in: Hennemann/Sattler (Hrsg.), Immaterialgüter und Digitalisierung, S. 89, 99; *Schack* in FS Wandtke, S. 9, 12.

448 *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261, 268. Zudem wird gegen die Anreizthese in Bezug auf Datenbanken die Argumentation herangezogen, dass die Datenbanken gerade nicht hergestellt werden, um diese wirtschaftlich zu verwerten, sondern sie häufig nur ein Nebenprodukt der eigentlichen wirtschaftlichen Betätigung darstellen, *C. König* in: Hennemann/Sattler (Hrsg.), Immaterialgüter und Digitalisierung, S. 89, 100; EuGH, ECLI:EU:C:2013:850 – *Innoweb/Wegener* = GRUR 2014, S. 166; EuGH, ECLI:EU:C:2015:10 – *Ryanair* = GRUR 2015, S. 253 m. Anm. Czychowski; EuGH, ECLI:EU:C:2015:735 – *Esterbauer* = GRUR 2015, S. 1187.

449 InfoSoc-RL Erw. 4.

450 BT-Drs. 18/12329, S. 19.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

hang mit gesetzlichen Nutzungserlaubnissen in der DSM-RL sowie im Umsetzungsgesetz betont.[451]

Resultierend daraus, dass sich die strikte Anreizthese schon nicht in Bezug auf traditionelles Werkschaffen und erst recht nicht in Bezug auf Innovation halten lässt, stellt sich die Frage, wie es das Urheberrecht stattdessen vermag, Innovation und technologischen Fortschritt zu steuern. Abschließend gilt also, dass auch für Innovation aus urheberrechtlicher Perspektive nicht zwingend ein hohes Schutzniveau ausschlaggebend ist,[452] es sind vielmehr auch hier die Beschränkungen heranzuziehen. Für eine Förderung von Innovation muss also der richtige Zuschnitt zwischen Ausschließlichkeitsrechten und inhaltlichen Beschränkungen erreicht werden.[453] Dieser kann auf unterschiedlichen Wegen erreicht werden, die einerseits die Rechtssetzung und andererseits die Rechtsauslegung betreffen.

Gerade im Zuge der Urheberrechtsreformen der letzten 20 Jahre wurde umfassend darüber diskutiert, ob flexiblere und entwicklungsoffene Schrankenbestimmungen, etwa nach Vorbild des amerikanischen Fair-Use-Systems, geeigneter sind, Wissenschaft und Innovation zu regulieren. Intendiert war dabei, dass die Offenheit einer Generalklausel, d. h. eine nicht auf bestimmt genannte Zwecke beschränkte Erlaubnisnorm, der *„legal lag"*, d. h. dem Hinterherhinken des Rechts hinter der sich schneller fortentwickelnden Technik entgegenwirken könnte, die urheberzentrierte Urheberrechtssysteme seit jeher prägt.[454] In Bezug auf diese Diskussion wurde insbesondere im Vorfeld der Reform, die schließlich im UrhWissG mündete, die sog. allgemeine Bildungs- und Wissenschaftsschranke in unterschiedlicher Reichweite diskutiert.[455]

---

451 DSM-RL Erw. 2, 5, 8.

452 *C. König* in: Hennemann/Sattler (Hrsg.), Immaterialgüter und Digitalisierung, S. 89, 99.

453 *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 14; ausführlich *Leistner*, ZGE 2009, S. 403, 406 ff.; *Leistner/Hansen*, GRUR 2008, S. 479, 483 f.

454 *Hofmann*, ZGE 2016, S. 482, 484; wohl auch *Hilty* in FS Schulze, S. 127, 132 f.; allgemein zum Technikrecht mit Verweis auf die Dynamik der Technik *Vieweg*, JuS 1993, S. 894, 896; *Vieweg* in FG Lukes, S. 199, 208 f.

455 Vgl. dazu die umfassende Untersuchung von *de la Durantaye,* Allgemeine Bildungs- und Wissenschaftsschranke, S. 207 ff., die schließlich eine Generalklausel mit Regelbeispielen vorschlug; ähnlich äußert sich auch *Leistner*, ZGE 2009, S. 403, 434; *Leistner/Hansen*, GRUR 2008, S. 479, 486 f., indem sie sich für einen Mittelweg zwischen richterlicher Rechtsauslegung und gesetzgeberischen Vorgaben aussprechen in *Leistner,* Konsolidierung und Entwicklungsperspektive des Europäischen Urheberrechts, S. 33 f. betont er demgegenüber die Folgen von Rechtsunsicherheit, die die Vorteile einer jedenfalls teilweisen und

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

Abgesehen von der Schwierigkeit, dass Generalklauseln mit dem unionsrechtlichen Urheberrechtssystem, in dem Schranken nur zu bestimmten Zwecken zulässig sowie mit dem Urheberpersönlichkeitsrecht nur schwer zu vereinbaren sind,[456] liegt der Frage nach der Innovationsfreundlichkeit von Generalklauseln eine Abwägung von Technologieoffenheit mit Rechtssicherheit und Technologiespezifität zugrunde. Es stellt sich also die Frage, ob eine Rechtsnorm dann als innovations- und forschungsfreundlich zu qualifizieren ist, wenn sie Rechtssicherheit bietet, oder dann, wenn sie für technische Fortentwicklungen genügend Spielraum bietet.[457] Aufgrund der Unvorhersehbarkeit technischen Fortschritts schränken technologiespezifische Schrankenregelungen den Anwendungsbereich stets ein, Generalklauseln hingegen bergen Risiken.[458]

Im aktuellen Urheberrecht lassen sich sowohl technologieneutrale als auch technologiespezifische Regelungen identifizieren. Insbesondere die Verwertungsrechte, § 15 UrhG, sind weitestgehend technologieneutral ausgestaltet, für die Schranken gilt das teilweise.[459] Ein Beispiel für eine

---

zwingend vergütungspflichtigen Generalklauseln aufheben könnten; in *Leistner*, Konsolidierung und Entwicklungsperspektive des Europäischen Urheberrechts, S. 33 f. betont er demgegenüber die Folgen von Rechtsunsicherheit, die die Vorteile einer jedenfalls teilweisen und zwingend vergütungspflichtigen Generalklauseln aufheben könnten; vgl. auch *Schack*, ZUM 2016, S. 266 ff., der sich ebenfalls für eine „*mittlere Lösung*" aussprach.

456 *Schack*, ZUM 2016, S. 266, 270; *de la Durantaye*, Allgemeine Bildungs- und Wissenschaftsschranke, S. 194 f., 206 f.; *Reinbothe*, ZGE 2015, S. 145, 160; *Wandtke/R. König*, ZUM 2014, S. 921, 922 ff.; *Wandtke*, GRUR 2015, S. 221, 223; kritisch auch *T. Dreier*, GRUR Int. 2015, S. 648 ff.; *Hansen*, Warum Urheberrecht?, S. 409; ebenso zur Berner Konvention mit Verweis auf den Drei-Stufen-Test *Reinbothe/v. Lewinski*, The WIPO treaties on copyright Rn. 7.10.17 mit weiteren Nachweisen.

457 So *T. Dreier*, GRUR Int. 2015, S. 648, 652; *Spindler*, NJW 2014, S. 2550, 2550 f.; *T. Dreier/Leistner*, GRUR 2013, S. 881, 889; *Hansen*, Warum Urheberrecht?, S. 405 ff. merkt an, dass aus den zahlreichen Urheberrechtsreformen ein Bedürfnis nach mehr Flexibilität resultiert. Das dürfte sich angesichts dessen, dass seitdem nicht weniger Reformen erfolgen mussten, um das Urheberrecht aktuellen technischen Gegebenheiten und wirtschaftlichen Realitäten anzuapassen, bestätigt haben; ebenso der sog. *„Reda-Bericht"*, Entwurf eines Berichts über die Umsetzung der Richtlinie 2001/29/EC des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft.

458 *„Rechtssicherheit lässt sich immer nur auf Kosten der Flexibilität erreichen, und umgekehrt."*, *Schack*, ZUM 2016, S. 266, 273.

459 Bzgl. der Weiterleitungshandlung *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 56; allgemeiner *Schulze* in Dreier/Schulze, UrhG, 7. Aufl. 2022,

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

technologiespezifische Schrankenbestimmung ist hingegen die Schranke zugunsten von vorübergehenden Vervielfältigungen, § 44a UrhG.[460] Dasselbe gilt für das TDM, wie später ausführlicher erläutert wird.

Die Frage, ob das Recht technologiespezifisch oder technologieneutral ausgestaltet sein sollte, setzt sich auf Ebene der Rechtsauslegung fort, denn auch die Rechtsauslegung hat das Potenzial, Ausnahmen und Beschränkungen bzw. Schrankenbestimmungen zu erweitern und damit den Zuschnitt des Urheberrechts zu verändern. Die Rechtsprechung hat urheberrechtliche Kompetenzen in der Vergangenheit einige Male ausgeweitet. Nennenswert ist etwa die Entscheidung des BGH zum elektronischen Kopienversand aus dem Jahr 1999, die es Bibliotheken erlaubte, Befugnisse für ihre Nutzer wahrzunehmen.[461] Im Fall um *Eugen Ulmer* und die *TU Darmstadt* ermöglichte der EuGH die Kumulation urheberrechtlicher Schrankenbestimmungen, sofern im jeweiligen Fall die jeweiligen Anforderungen erfüllt sind.[462] In der *Used-Soft*-Rechtsprechung ebnete er außerdem den Weg, um unter gewissen Voraussetzungen gebrauchte Software weiter veräußern zu können.[463]

Ein Ansatz zur urheberrechtlichen Innovationsförderung auf Ebene der Schrankenauslegung ist die technikfreundliche Schrankenauslegung

---

§ 2 Rn. 199: „*Das Urheberrecht ist tendenziell technologieneutral und will alle Vorgänge sowohl zur Herstellung als auch zur Verwertung von Werken erfassen, seien sie analog, digital oder auch in anderer Technik*“, § 15 Rn. 10; *v. Ungern-Sternberg* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 15 Rn. 16; bzgl. dem Schutz von Lichtbildwerken *Loewenheim/Leistner* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 2 Rn. 208; BGH, Urt. v. 5.7.2001 - I ZR 335/98 – *Scanner* = GRUR 2002, S. 246, 247; BVerfG, Beschl. v. 30.8.2010 - 1 BvR 1631/08 – *Drucker und Plotter* = GRUR 2010, S. 999, 1002 Rn. 64. Gleichzeitig orientieren sich die Schranken in erster Linie am Stand der Technik zum Zeitpunkt ihrer Schaffung, obwohl auch erweiterte Auslegungen möglich sind, vgl. dazu *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Vor § 44a Rn. 7a; BGH, Urt. v. 16.01.1997 - I ZR 9/95 – *CB-Infobank I* = GRUR 1997, S. 459, 463; BGH, Urt. v. 18.05.1955 - I ZR 8/54 – *Magnettonbänder II* = GRUR 1955, S. 492, 496 f. sowie die bisherigen Ausführungen zur erweiterten Schrankenauslegung, S. 76 ff.

460 *Hofmann*, ZGE 2016, S. 482, 506; *Specht*, GRUR 2019, S. 253, 254; *Hilty* in FS Schulze, S. 127, F22 f. fordert eine weiterreichende Technologieneutralität der Schrankenregelungen, um Raum für technische Entwicklungen zu schaffen.

461 BGH, Urt. v. 25.2.1999 - I ZR 118/96 – *Kopienversanddienst* = GRUR 1999, 707; *Hofmann*, ZGE 2016, S. 482, 485 verweist außerdem auf § 53a UrhG a. F., der heute in § 60e Abs. 5 UrhG geregelt ist.

462 EuGH, ECLI:EU:C:2014:2196 – *Eugen Ulmer* = GRUR 2014, 1078.

463 EuGH, ECLI:EU:C:2012:407 – *Used Soft I* = GRUR 2012, 904; BGH, Urt. v. 17.7.2013 - I ZR 129/08 – *Used Soft II* = GRUR 2014, S. 264; BGH, Urt. v. 11.12.2014 – I ZR 8/13 – *Used Soft III* = GRUR 2015, S. 772.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

(„*more technological approach*"), d. h. eine „*für die Bedürfnisse der technischen Umwelt aufgeschlossene Analyse des Immaterialgüterrechts*".[464] Angestrebt wird also eine Rechtsbeeinflussung durch Technik.[465] Ein wichtiges Argument für die Vornahme einer grundsätzlich technologiefreundlichen Auslegung ist, dass technologiespezifische Sachverhalte längst Einklang in das Urheberrecht gefunden haben. Eine Technikbezogenheit des Urheberrechts ist auch nicht erst durch die Digitalisierung entstanden, auch zuvor wurden neue Methoden bewusst adressiert und bestimmte Verwertungsrechte entsprechend ausgelegt.[466] Zweifel an der technologiefreundlichen Rechtsauslegung bestehen jedoch aus systematischer Perspektive, denn das Urheberrecht ist ursprünglich nicht als Technikrecht geschaffen.[467] Als Gegenmodell wurde eine technologieneutrale Auslegung vorgeschlagen.[468]

---

464 Dennoch solle das Recht, nicht aber die Technik, entscheiden, welche Erwartungshaltung schutzwürdig ist und welche nicht, *Grünberger*, ZUM 2016, S. 905, 911; *Podszun/Grünberger*, ZGE 2014, S. 269, 270; *Specht*, GRUR 2019, S. 253, 255. Der Ansatz wird vor allem auf die Rechtsprechung zum Linking und Framing gestützt, *Hofmann*, ZGE 2016, S. 482, 486 f.; *Grünberger*, ZUM 2015, S. 273, 276: „*Ein Berechtigter, der ein urheberrechtlich geschütztes Werk ohne technische Schutzmaßnahmen im Internet öffentlich zugänglich macht, ermöglicht dadurch bereits selbst die Nutzungen, die ein Abrufender vornehmen kann. Es ist seine Entscheidung, ob er das Werk trotz der Möglichkeit, dass nach Abruf auch rechtswidrige Nutzungen vorgenommen werden, weiter zum Abruf bereithält.*", BGH, Urt. v. 17. 7. 2003 - I ZR 259/00 – *Paperboy* = GRUR 2003, S. 958, 961; ebenso BGH, Urt. v. 29. 4. 2010 - I ZR 39/08 – *Session-ID* = GRUR 2011, S. 56, 58 Rn. 27: „*Macht ein Berechtigter ein urheberrechtlich geschütztes Werk ohne technische Schutzmaßnahmen im Internet öffentlich zugänglich, haftet der Linksetzende nicht als Störer oder Täter, weil der Berechtigte dadurch bereits selbst den Nutzern die Vervielfältigung ermöglicht und den Zugang eröffnet hat und der Linksetzende die ohnehin mögliche Vervielfältigung und den ohnehin eröffneten Zugang lediglich erleichtert*".

465 *Hofmann*, ZGE 2016, S. 482, 487; *Specht*, GRUR 2019, S. 253, 255.

466 *Hofmann*, ZGE 2016, S. 482, 484; *Specht*, GRUR 2019, S. 253, 253 f.; *Dreier* in Dreier/Schulze, UrhG, 7. Aufl. 2022, Einleitung Rn. 11; BGH, Urt. v. 18.05.1955 - I ZR 8/54 – *Grundig-Reporter* = GRUR 1955, S. 492.

467 *M. Becker*, ZGE 2016, S. 239, 267 ff.; das räumt auch *Hofmann*, ZGE 2016, S. 482, 489, 502, 505 ein.

468 *Hofmann*, ZGE 2016, S. 482, 483, 511 mahnt aber einschränkend die Notwendigkeit eines gesunden Realismus an, ihm zufolge handelt es sich bei Technologieneutralität um ein „*rechtspolitisches Postulat*"; nach *B. Raue*, ZGE 2017, S. 514, 515 ist das Recht „*situationsgebunden auszulegen*", weswegen er dafür plädiert, es auf den sozialen und wirtschaftlichen Sinngehalt und nicht auf die technische Ausgestaltung ankommen zu lassen; *Heinzmann* in: Taeger (Hrsg.), Internet der Dinge, S. 599, 609 ff., die an eine wirtschaftliche Relevanz sowie an eine öffentliche Zugänglichmachung anknüpfen möchte; vgl. auch *v. Ungern-Sternberg* in Schricker/Loewenheim, UrhG, 6. Aufl. 2020, § 15 Rn. 15 f., 222. Auch

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

D. Stellungnahme

Eine gewisse Innovationsfreundlichkeit ist dem deutschen Urheberrechtssystem bereits immanent, denn wissenschaftliche Erkenntnisse sind urheberrechtlich nicht geschützt, außerdem sind gesetzliche Schrankenbestimmungen aus der grundgesetzlich geschützten Eigentumsgarantie ausgenommen. Die Annahme, Werkschaffen sei durch ein hohes Schutzniveau sowie Restriktion der Nutzungserlaubnisse zu motivieren, ist bereits in ihrem Ausgangspunkt unzutreffend. Dasselbe gilt für wissenschaftliche Tätigkeit und Innovation. Das Urheberrecht kann allerdings mittels des richtigen Zuschnitts von Ausschließlichkeits- und Nutzungsrechten als Innovationsbeschleuniger fungieren.

Den dargestellten Ansätzen zur Innovationsförderung ist differenzierend zu begegnen. Wie bereits dargestellt wurde, ist nicht nur die Schaffung von Werken Sinn und Zweck des Urheberrechts, denn gerade die Schrankenbestimmungen definieren erst den Schutzbereich des Urheberrechts. Insgesamt sollte sich die Rechtsauslegung jeweils am Telos der betreffenden Norm orientieren, die technologieneutral gemeint sein kann, es aber nicht zwingend sein muss. Allgemeine Grundsätze lassen sich jedoch nicht bilden. Vorzuziehen ist deswegen eine Kombination der verschiedenen Lösungsansätze. Es erscheint zielführend, eine urheberrechtliche Innovationsförderung durch ein Ineinandergreifen von Rechtsfortbildung sowie Ausnutzung bzw. Auslegung bestehender Schrankenbestimmungen anzustreben. Außerdem ist es als relevant einzustufen, bestehende Erlaubnisnormen auch tatsächlich zu nutzen sowie gerichtlich auf eine erweiterte Auslegung hinzuwirken, denn nur durch eine Rechtswahrnehmung kann schließlich auch die notwendige Rechtssicherheit erlangt werden.[469]

Die Frage, auf welche Art und Weise das Urheberrecht Anreize für innovative Tätigkeiten setzen kann, ist deswegen nicht generell, sondern

---

dieser Auslegungsansatz kann sich auf höchstrichterliche Rechtsprechung stützen, mit Bezug auf u. a. EuGH, ECLI:EU:C:2006:764 – *SGAE/Rafael Hoteles SL* = GRUR 2007, 225, S. 227 Rn. 46; EuGH, ECLI:EU:C:2011:631 – *Football Association Premier League u. Murphy* = GRUR 2012, S. 156, 165 Rn. 193; EuGH, ECLI:EU:C:2009:465 – *Infopaq* = GRUR 2009, S. 1041, 1044 Rn. 41 ff.; vgl. dazu auch *Jani/Leenen*, GRUR 2014, S. 362, 362; *v. Ungern-Sternberg*, GRUR 2012, S. 1198, 1205 bescheinigt dem EuGH eine funktionsbezogene Auslegungsmethode; ebenso *Leistner*, EuZW 2016, S. 166, 168.

469 Diesbezüglich sei auf die strategische Prozessführung hingewiesen, die von *Reda/Binder*, RuZ 2020, S. 176 ff. vorgestellt wird, die allerdings in erster Linie das Ziel des erweiterten Informationszugangs verfolgen.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

individuell beantworten. Deswegen gilt es, entsprechend der praktischen Bedarfe gezielte Rechtsanpassungen vorzunehmen. Dazu muss differenziert werden, welche Rechte betroffen sind, welchen Personengruppen die betreffenden Rechteinhaber angehören und welche Nutzungshandlungen verboten oder erlaubt sein sollen.[470] Nur unter genauer Betrachtung individueller Ausschließlichkeits- und Nutzungsrechte lässt sich ein Zuschnitt des Urheberrechts, der zugleich Anreize für Geistesschaffen und -nutzen setzt, erreichen.[471] Zielführend erscheint es deswegen, empirische Daten zu erheben und auszuwerten.[472] Generelle Auslegungstendenzen sind hingegen weder zielführend noch dogmatisch mit dem Urheberrecht zu vereinbaren.

§ 6 Ergebnis des ersten Teils

Das TDM ist ein technisches Verfahren, das zu Zwecken der wissenschaftlichen Forschung zur Erzielung unmittelbarer Erkenntnisse genutzt werden kann, zu dem aber auch Methoden des Forschungsfelds der Künstlichen Intelligenz zählen, so basiert etwa das Training im Bereich des maschinellen Lernens auf dem technischen Verfahren des TDMs. Insgesamt bezieht sich der hier genutzte Begriff des TDMs auf den Gesamtprozess von der Datengewinnung über die Vorverarbeitung bis hin zur eigentlichen algorithmusbasierten Analyse.

Das Urheberrecht reguliert die wissenschaftliche Betätigung auf eine besondere Art und Weise. Durch die genaue Differenzierung, welche Elemente wissenschaftlicher Erzeugnisse aus urheberrechtlicher Perspektive schutzfähig sind, soll Wissen möglichst frei kursieren. Ebenso wichtig sind allerdings geeignete Schrankenbestimmungen, die den wissenschaftli-

---

470 Dieses Ergebnis erfährt breite Akzeptanz in der Literatur, vgl. *Ohly* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 279, 297; *Peukert* in: Eifert/Hoffmann-Riem (Hrsg.), Geistiges Eigentum und Innovation, S. 299, 302; *T. Dreier* in: Klumpp/Kubicek/Roßnagel/Schulz (Hrsg.), Medien, Ordnung und Innovation, S. 261 ff.; *Kerber*, ZGE 2013, S. 245, 271 kritisiert in diesem Kontext allerdings, dass eine zunehmende Spezifizierung und Differenzierung eine zu große Rechtsunsicherheit und damit auch wirtschaftliche Nachteile beinhalten würde.

471 Ähnlich *Kerber* in: Grünberger/Leible (Hrsg.), Die Kollision von Urheberrecht und Nutzerverhalten, S. 31, 59.

472 In diesem Zusammenhang spricht *Leistner*, ZGE 2009, S. 403, 412 von einer *„Rationalisierung ökonomischer Folgenüberlegungen"*.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb

chen Urhebern den notwendigen Freiraum einräumen. Bei der Auslegung der Schrankenbestimmungen kann die Wissenschaftsfreiheit erweiternd berücksichtigt werden, sofern grundlegende Grundrechtswertungen ausgehebelt werden. Das sind im besonderen Maße die wissenschaftlichen Eigengesetzlichkeiten, die eine Arbeit erst als wissenschaftlich qualifizieren.

Bei der Analyse und Verbesserung eines technologieadäquaten Rechtsrahmens zugunsten des TDMs sowohl die Charakteristika des Wissenschaftsurheberrechts als auch die Besonderheiten der urheberrechtlichen Innovationssteuerung zu beachten. Zur Regulierung von Innovation durch das Urheberrecht ist die Ausgestaltung von Beschränkungen des eigentlichen Urheberrechts als Recht des Urhebers essenziell. Anreize können über ein angemessenes Verhältnis von Technologieneutralität und -spezifität, d. h. einen Ausgleich zwischen Rechtssicherheit und Entwicklungsoffenheit, erreicht werden.

https://doi.org/10.5771/9783748935360-29, am 31.07.2024, 14:28:49
Open Access – (cc) BY – https://www.nomos-elibrary.de/agb