

FELICITAS KLEINKOPF

TEXT- UND DATA-MINING IN DEN DIGITALEN GEISTESWISSENSCHAFTEN

Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen“

Text- und Data-Mining in den digitalen Geisteswissenschaften

Gutachten zu den urheberrechtlichen
Rahmenbedingungen im Rahmen
des Forschungsprojektes „XSample –
Text- und Data-Mining auf
geschützten Werken durch Auszüge
transparent erschließen“

Felicitas Kleinkopf

]u[

ubiquity press
London

Published by
Ubiquity Press Ltd.
Unit 3N, 6 Osborn Street
London E1 6TD
www.ubiquitypress.com

Text © Felicitas Kleinkopf 2023
First published 2023
Universität Stuttgart / University of Stuttgart
Series: Stuttgart University Reports, 1

Supporting Agency
Ministerium für Wissenschaft, Forschung und Kunst
Baden-Württemberg (MWK)
Bundesministerium für Bildung und Forschung (BMBF)
Grant Number 6TOA022B (TU9_Monos)

Cover design by KIT Scientific Publishing
Cover image by Penny Travlou

Print and digital versions typeset by Siliconchips Services Ltd.

ISBN (Paperback): 978-1-914481-28-4
ISBN (PDF): 978-1-914481-29-1
ISBN (EPUB): 978-1-914481-30-7
ISBN (Mobi): 978-1-914481-31-4
DOI: <https://doi.org/10.5334/bcu>

This work is licensed under the Creative Commons Attribution 4.0 International License (unless stated otherwise within the content of the work). To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, USA. This license allows for copying any part of the work for personal and commercial use, providing author attribution is clearly stated.

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see <http://www.ubiquitypress.com/>

Suggested citation:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. London: Ubiquity Press.
DOI: <https://doi.org/10.5334/bcu>. License: CC BY 4.0

To read the free, open access version of this book online, visit <https://doi.org/10.5334/bcu> or scan this QR code with your mobile device:



Autorin

Felicitas Kleinkopf Zentrum für Angewandte Rechtswissenschaft (ZAR),
Institut für Informations- und Wirtschaftsrecht (IIWR),
Karlsruher Institut für Technologie

Beteiligte

Melanie Andresen Institut für Maschinelle Sprachverarbeitung, Universität
Stuttgart
Markus Gärtner Institut für Maschinelle Sprachverarbeitung, Universität
Stuttgart
Sibylle Hermann Universitätsbibliothek Stuttgart
Janina Jacke Institut für Maschinelle Sprachverarbeitung, Universität
Stuttgart
Axel Pichler Institut für Maschinelle Sprachverarbeitung, Universität
Stuttgart

Projektverantwortliche

Prof. Dr. Jonas Kuhn, Institut für Maschinelle Sprachverarbeitung, Universität
Stuttgart

Dr. Helge Steenweg, Universitätsbibliothek Stuttgart

Prof. Dr. iur. Thomas Dreier M.C.J., Institut für Informations- und Wirtschaftsrecht (IIWR), Zentrum für Angewandte Rechtswissenschaft (ZAR), Karlsruher Institut für Technologie

Dieses Gutachten zum rechtlichen Umgang mit der digitalen Forschungsmethode des Text- und Data-Mining wurde mit größter Sorgfalt und nach der guten wissenschaftlichen Praxis erstellt. Die vermittelten Erkenntnisse lassen sich allerdings nicht ohne weiteres auf andere Forschungsprojekte übertragen. Das Gutachten stellt daher keine Rechtsberatung, sondern eine Einschätzung eines neuen Forschungsbereichs dar. Bitte wenden Sie sich bei Rechtsfragen und experimentellen Forschungsprojekten mit urheberrechtlich geschützten Werken an einen zugelassenen Rechtsanwalt oder das Justizariat Ihrer Forschungseinrichtung. Das Projektteam freut sich über wissenschaftlichen Diskurs per E-Mail. Für Schäden oder sonstige Folgen eines rechtswidrigen Umgangs im Bereich der Text- und Data-Mining-Forschung übernimmt das Projektteam keine Haftung.

Weitere Informationen zum XSample-Projekt sowie dem Rahmenprojekt „BW-BigDIWA – Wissenschaftliche Bibliotheken gestalten den Digitalen Wandel“ finden Sie unter:

<https://www.izus.uni-stuttgart.de/fokus/fdm-projekte/xsample/>

<https://bw-bigdiwa.bib.uni-mannheim.de/projekte/xsample-stuttgart/>

Stand: 01.02.2022

Inhaltsverzeichnis

Abbildung	vii
Abkürzungsverzeichnis	ix
1. Einleitung	1
2. Grundlagen	5
2.1 Einführung in das Urheberrecht	5
2.2 Gesetzesentwicklung	7
3. Das praktizierte Text- und Data-Mining in den digitalen Geisteswissenschaften	9
3.1 Zusammenfassung des Forschungsprozesses	9
3.2 Status quo in den Digital Humanities	10
3.3 Anwendungsfälle	11
3.3.1 <i>Erster Anwendungsfall: Unzuverlässiges Erzählen</i>	11
3.3.2 <i>Zweiter Anwendungsfall: Forschungsliteratur und Wissenschaftssprache</i>	11
4. Rechtliche Beurteilung	13
4.1 Berechtigte	13
4.2 Schutz des Ursprungsmaterials	13
4.3 Zugriff auf das Ursprungsmaterial: Rechtmäßiger Zugriff zwischen Verträgen und technischen Schutzmaßnahmen	15
4.3.1 <i>Vertragsfestigkeit der TDM-Schranke</i>	15
4.3.2 <i>Durchsetzbarkeit der urheberrechtlichen Schranke gegenüber technischen Schutzmaßnahmen, §§ 60g, 95b UrhG</i>	16
4.3.3 <i>Andere Umgehungsmöglichkeiten</i>	16
4.4 Digitalisierung analoger Quellen	17
4.4.1 <i>Korrekte Digitalisierung</i>	17
4.4.2 <i>Fehlerhafte Digitalisierungen</i>	19
4.5 Texterkennung mithilfe von OCR	21
4.5.1 <i>Korrekte Texterkennung</i>	21
4.5.2 <i>Fehlerhafte Texterkennungen</i>	22
4.5.3 <i>Händische Übertragungen</i>	22

4.6 Segmentierung	23
4.6.1 <i>Vervielfältigungen und Bearbeitungen, §§ 16 Abs. 1, 23 Abs. 1 UrhG</i>	23
4.6.2 <i>Änderungen, § 62 UrhG</i>	25
4.7 Bereinigung der Texte	25
4.8 Annotationen	26
4.8.1 <i>Bearbeitungen, § 23 Abs. 1 UrhG</i>	26
4.8.2 <i>Änderungen, § 62 UrhG</i>	27
4.9 Folgeprobleme bei der Weiterverwendung veränderter oder bearbeiteter Kopien	28
4.9.1 <i>Umgang mit dem Bearbeitungsrecht, § 23 UrhG</i>	28
4.9.2 <i>Umgang mit dem Änderungsverbot, § 62 UrhG</i>	28
4.10 Pflicht zur Quellenangabe, § 63 UrhG	30
4.11 Das Korpus	31
4.11.1 <i>Schutz der Korpora</i>	31
4.11.2 <i>Rechteinhaberschaft</i>	34
4.11.3 <i>Rechtsverletzungen durch Integrierung der aufbereiteten Werke in ein Forschungskorpus</i>	35
4.12 Gemeinsame Forschungsarbeiten	36
4.13 Rechtliche Bewertung der TDM-Analyse im engeren Sinne	37
4.14 Vergütung	38
4.15 Publikation der Forschungsergebnisse	38
5. Umgang mit den Korpora nach Abschluss der Forschungsarbeiten	41
5.1 Aufbewahrung der Korpora	41
5.2 Überprüfbarkeit der wissenschaftlichen Forschung	43
5.3 Nachnutzung von TDM-Korpora	44
5.3.1 <i>Terminal-Kopien, § 60e Abs. 4 UrhG</i>	45
5.3.2 <i>Nicht-kommerzielle, wissenschaftliche Forschung, § 60c UrhG</i>	46
5.3.3 <i>Auszugskonzept</i>	49
6. Fazit	57
Anhang: Historie der relevanten Rechtsnormen	59
Literaturverzeichnis	65

Abbildung

1. Segmentierung	24
2. Annotationen	26
3. Auszugskonzept	50
4. Nutzbarkeit kurzer Werke aus Sammelwerken	55

Abkürzungsverzeichnis

a. F.	Alte Fassung
Erw.	Erwägungsgrund
UrhWissG	Gesetz zur Angleichung des Urheberrechts an die aktuellen Erfordernisse der Wissensgesellschaft
mwN	Mit weiteren Nachweisen
n. F.	Neue Fassung
OCR	Optical Character Recognition
Rn.	Randnummer
DSM-RL	Richtlinie (EU) 2019/790 des Europäischen Parlaments und des Rates vom 17. April 2019 über das Urheberrecht und die verwandten Schutzrechte im digitalen Binnenmarkt und zur Änderung der Richtlinien 96/9/EG und 2001/29/EG
InfoSoc-RL	Richtlinie 2001/29/EG des Europäischen Parlaments und des Rates vom 22. Mai 2001 zur Harmonisierung bestimmter Aspekte des Urheberrechts und der verwandten Schutzrechte in der Informationsgesellschaft
Datenbank-RL	Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken
TDM	Text- und Data-Mining
UrhG	Urheberrechtsgesetz

Einleitung

Unter der Bezeichnung „*Text- und Data-Mining*“ (im Folgenden: TDM) wird ein digitales Analyse-verfahren verstanden, mithilfe dessen große Datenmengen automatisiert auf bestimmte, dem Menschen ansonsten nicht¹ zugängliche Informationen, Korrelationen und Zusammenhänge untersucht werden können. In diesem Zuge wird aus verschiedensten Ursprungsmaterialien, die umfassend vorverarbeitet werden, ein sog. Korpus erstellt, das dann analysiert wird. Gerade in den digitalen Geisteswissenschaften (engl. Digital Humanities) erfreuen sich TDM-Methoden einer immer größer werdenden Beliebtheit.

Lange war umstritten, ob das TDM-Verfahren überhaupt urheberrechtlich relevant ist. Die Europäische Union sowie der deutsche Gesetzgeber haben jedoch klargestellt, dass das TDM-Verfahren jedenfalls dann urheberrechtlich nicht relevant ist, wenn die Textanalyse ohne Vervielfältigungen oder öffentliche Zugänglichmachung, d. h. Verwertungen urheberrechtlicher Schutzgegenstände, auskommt, und die maschinelle Analyse als solche aus urheberrechtlicher Perspektive freigestellt ist.² Dasselbe gilt selbstverständlich auch, wenn gemeinfreies Gut untersucht wird, das entweder nie urheberrechtlich geschützt war oder dessen Schutzfrist abgelaufen ist.³ Die **Vorverarbeitung** der Forschungsdaten, d. h. die Erstellung des Korpus, erfolgt aber regelmäßig

¹ Oder nur sehr erschwert, vgl. Jannidis/Kohle/Rehbein, Digital Humanities, S. 3.

² BT-Drs. 18/12329, S. 40; DSM-RL Erw. 9; Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 3; Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, § 60d Rn. 4; Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60d Rn. 1; Raue, ZUM 2019, S. 684, 685.; Geiger/Frosio/Bulanyenko, IIC 2018, S. 814, 817; Wandtke, NJW 2019, S. 1841, 1841.

³ Die urheberrechtliche Schutzfrist beginnt mit der Schöpfung des Werks und endet 70 Jahre nach dem Tod des Urhebers, § 64 UrhG.

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 1–3. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.a>. License: CC BY 4.0

dergestalt, dass **Vervielfältigungen, § 16 Abs. 1 UrhG, Bearbeitungen, § 23 UrhG** erstellt oder das Werk bzw. Kopien des Werks **öffentlich zugänglich gemacht** werden, **§ 19a UrhG**, um das Forschungsmaterial für die eigentliche Analyse zu strukturieren und aufzubereiten. Diese Handlungen bedürfen einer rechtlichen Erlaubnis, einer sog. **urheberrechtlichen Schranke** oder einer **rechtsgeschäftlichen Erlaubnis**. Zur Vermeidung von Konflikten oder Rechtsunsicherheiten langfristig auf die Forschung an urheberrechtlichen Schutzgegenständen zu verzichten, ist hingegen keine Option: Der Forschung bliebe dadurch ein großes Potenzial an neuen Erkenntnissen vorenthalten.⁴ Aufgrund der Größe der zu analysierenden Datensätze ist es auch nicht praktikabel, im Einzelnen die Zustimmung der Rechteinhaber einzuholen bzw. Lizenzverträge abzuschließen.⁵

Trotz der bereits seit dem UrhWissG, in Kraft seit dem 1.3.2018, in Deutschland geltenden **Erlaubnis, mit TDM an urheberrechtlich geschützten Dokumenten zu forschen – § 60d UrhG** – die zum 7.6.2021 auf Grundlage der europäischen DSM-RL umfassend reformiert wurde, wird TDM an urheberrechtlichen Schutzgegenständen bislang nicht regelmäßig praktiziert.⁶ Das ist unter anderem auf **Rechtsunsicherheiten** zurückzuführen, die in Bezug auf das TDM an urheberrechtlichen Schutzgegenständen galt und auch durch die Reformen nicht vollständig behoben werden konnten. Insbesondere in den digitalen Geisteswissenschaften bestehen diese Rechtsunsicherheiten betreffend der Frage, ob sämtliche Veränderungen der urheberrechtlichen Werke, die innerhalb der Vorverarbeitung der Forschungsdaten für das TDM erfolgen, tatsächlich von der Erlaubnisnorm gedeckt werden. Ebenso ist die Frage nach dem Umgang mit den Forschungsdaten nach Abschluss der Forschungsarbeiten unklar, d. h. inwieweit die Archivierung sowie wissenschaftliche Überprüfung möglich sind und ob die Forschungsdaten nach Projektabschluss weiterverwendet werden dürfen. Infolge dieser Unsicherheiten wird oft auf die Weiterverwendung der Korpora verzichtet, wenn denn überhaupt eine Forschung an urheberrechtlich geschützten Materialien erfolgt.

Das Projekt **XSample** verfolgt das Ziel, die Forschung an urheberrechtlich geschützten Gütern in den digitalen Geisteswissenschaften zu etablieren,

⁴ Aus diesem Grund kommt etwa die gelegentlich vorgeschlagene Filterung urheberrechtlich geschützten Materials aus zu beforschenden Datensätzen nicht in Betracht. Zudem ließe sich nur in den seltensten Fällen mit absoluter Sicherheit feststellen, dass kein urheberrechtlich oder durch verwandte Schutzrechte geschütztes Material mehr in den Korpora enthalten sind, vgl. dazu Bunk in Smart-Data-Begleitforschung, Forschungszentrum Informatik (Hrsg.), Daten als Wirtschaftsgut, S. 30, 32.

⁵ Raue, CR 2017, S. 656, 661; Spindler, GRUR 2016, S. 1112, 1118; Andresen et al., ZfdG im Erscheinen 2022; Schöch et al., ZfdG 2020 Rn. 9.

⁶ Vgl. auch Schöch et al., ZfdG 2020 Rn. 1; Raue/Schöch, RuZ 2020, S. 118, 118 f.

indem ein denkbarer Forschungsprozess transparent dokumentiert und analysiert wird. Anhand von realen Anwendungsfällen wird herausgearbeitet, welche Methoden des TDM in fremde Urheberrechte eingreifen und zum anderen, ob diese Eingriffe durch die betreffenden urheberrechtlichen Schranken gerechtfertigt sind (s. dazu Abschnitt 4). Insofern wird im folgenden Gutachten untersucht, ob die Rechtslage den Forschungsprozess adäquat abdeckt. Des Weiteren soll der Umgang mit den Korpora nach Forschungsabschluss beleuchtet sowie die Frage nach der Nachnutzbarkeit nach Unionsrecht und deutschem Urheberrecht beantwortet werden (s. dazu Abschnitt 5).

KAPITEL 2

Grundlagen

Untersuchungsgegenstand des folgenden Gutachtens ist die Bewertung von TDM-Forschung nach dem Urheberrecht.⁷

2.1 Einführung in das Urheberrecht

Der Schutzgegenstand des Urheberrechts ist das **Werk**, § 1 UrhG. § 2 Abs. 1 UrhG enthält dazu eine (nicht abschließende) Aufzählung an Werktypen, wie u. a. **Sprachwerke** (Schriftwerke, Reden und Computerprogramme), **Werke der Musik**, **Lichtbildwerke** oder **Werke der bildenden Künste**. Daneben können Sammlungen unabhängigen Elemente als **Sammel- oder Datenbankwerke** geschützt sein, § 4 UrhG. Die entscheidende Voraussetzung ist allerdings in § 2 Abs. 2 UrhG normiert: Ein urheberrechtliches Werk kann nur eine **persönliche geistige Schöpfung** sein. Der Werkbegriff ist in seinem Grundsatz nationalstaatlich definiert und findet im EU-Recht nur vereinzelt Erwähnung.⁸ Dennoch ist inzwischen ein stärkerer Einfluss der Rechtsprechung des EuGH auf den Werkbegriff zu verzeichnen.⁹ Dort bildet das Kriterium der **eigenen**

⁷ Grundsätzlich kann zwar auch das Datenschutzrecht berührt sein, wenn personenbezogene Daten betroffen sind, vgl. Art. 1 Abs. 1, 2 Abs. 1 DS-GVO. Verwiesen sei dazu aber etwa auf Spindler, GRUR 2016, S. 1112, 1116 f; Spindler, ZGE 2018, S. 273, 297 f.

⁸ Vgl. Art. 1 Abs. 3 Computerprogramm-RL, Richtlinie 2009/24/EG, Art. 3 Abs. 1 Datenbank-RL, Richtlinie 96/9/EG sowie Art. 6 Schutzdauer-RL, Richtlinie 2006/116/EG.

⁹ Handig, GRUR Int. 2012, S. 973, 974; Loewenheim/Leistner in Loewenheim UrhR-HdB, § 6 Rn. 6.; EuGH, ECLI:EU:C:2009:465, Infopaq = GRUR

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 5–8. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.b>. License: CC BY 4.0

geistigen Schöpfung die Voraussetzung urheberrechtlichen Werkschutzes, das niedrigere Anforderungen genügen lässt.¹⁰ Die urheberrechtlichen Schutzvoraussetzungen seien im Folgenden in gebotener Kürze umrissen: Unter dem Kriterium der *persönlichen* Schöpfung ist zu verstehen, dass das Werk vom Menschen geschaffen wurde.¹¹ Das Kriterium der *geistigen* Schöpfung beinhaltet, dass das Immaterialgut, das im Werkstück konkretisiert ist, eine geistige Anregung auf den Betrachter ausübt.¹² Erforderlich ist, dass der menschliche Geist zum Ausdruck kommt.¹³ Das i. d. R. entscheidende Kriterium ist die *Individualität* bzw. die *Gestaltungs- oder Schöpfungshöhe*. Diese Schöpfungshöhe setzt voraus, dass die Persönlichkeit des Urhebers im Werk verkörpert sein muss, was das Werk von anderen Werken abhebt.¹⁴ Bei literarischen Werken stehen die Schutzvoraussetzungen zumeist nicht in Zweifel. Aber auch wissenschaftliche Texte oder Darstellungen können urheberrechtlichen Werkschutz für sich beanspruchen. Dort ist zu beachten, dass nicht die wissenschaftlichen Thesen oder Theorien geschützt sind, sondern die Art und Weise, wie diese dem Rezipienten vermittelt werden.¹⁵ Insgesamt sind die Anforderungen jedoch nicht allzu hoch angesiedelt, es können auch bereits Werkteile von nur elf Wörtern eine eigene geistige Schöpfung zum Ausdruck bringen.¹⁶

2009, 1041, S. 1044 Rn. 35; EuGH, Urt. v. 4. 10. 2011 – C-403, 429/08, ECLI:EU:C:2011:631, Wettbewerbsverstoß durch territoriale Exklusivitätsvereinbarungen bei Fußball-Übertragungen = GRUR 2012, 156, S. 165 Rn. 188; EuGH, ECLI:EU:C:2012:115, Football Dacato/Yahoo = GRUR 2012, 386, S. 165 Rn. 188.

¹⁰ Der Begriff wurde zunächst nur in Bezug auf Datenbankwerke Computerprogramme und Lichtbildwerke (also Fotografien) geprägt, ob er auf alle Werkarten übertragen werden kann, ist allerdings umstritten, vgl. Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 22. Doch auch die Übernahme eines Textausschnittes von elf Wörtern betrachtete der EuGH als Eingriff in das ausschließliche Recht des Urhebers, sein Werk zu vervielfältigen, EuGH, ECLI:EU:C:2009:465, Infopaq = GRUR 2009, 1041, S. 1041 Rn. 51.

¹¹ Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 38; Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 8.

¹² Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 12.

¹³ Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 45.

¹⁴ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 18; Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 50.; Bullinger in Wandtke/Bullinger, UrhG, § 2 Rn. 21.

¹⁵ Diese Thematik war in der Vergangenheit stark umstritten, dazu Loewenheim/Leistner in Loewenheim UrhR-HdB, § 7 Rn. 9; Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 86.

¹⁶ Wiebe in Spindler/Schuster, Recht der elektronischen Medien, § 2 UrhG Rn. 3; Wiebe in Spindler/Schuster, Recht der elektronischen Medien, § 2

Ideell basiert das Urheberrecht auf zwei Säulen, einerseits dem Urheberpersönlichkeitsrecht, d. h. die persönlichkeitsrechtlichen, und andererseits den vermögensrechtlichen Bestandteilen, die als untrennbare Einheit eingeordnet werden (sog. monistische Theorie).¹⁷ Aus der persönlichkeitsrechtlichen Seite resultieren Rechte des Urhebers, sein Werk zu veröffentlichen (§ 12 UrhG), als Urheber anerkannt zu werden (§ 13 UrhG) sowie das Recht, bestimmte Entstellungen oder andere Beeinträchtigungen seines Werks zu verbieten (§ 14 UrhG). Aus der vermögensrechtlichen Seite resultiert u. a. das grundsätzliche Recht des Urhebers, sein Werk zu vervielfältigen (§ 16 UrhG) und es öffentlich zugänglich zu machen (§ 19a UrhG).

Neben dem Werkschutz sieht das UrhG auch **verwandte Schutzrechte** bzw. **Leistungsschutzrechte** vor, z. B. das **Datenbankherstellerrecht**, §§ 87 a ff. UrhG. Der Schutzgrund des Datenbankherstellerrechts ist vorwiegend wirtschaftspolitischer Natur.¹⁸ Daneben können Fotografien, die nicht hinreichend individuell sind, um als urheberrechtliches Werk zu gelten, als **einfache Lichtbilder** geschützt sein, § 72 UrhG.

Neben denjenigen Personenkreisen, die Werke oder verwandte Schutzgegenstände erschaffen haben und ein Interesse an dem Schutz ihrer Erzeugnisse und an der Beteiligung derer Verwertung haben, haben auch Dritte berechnigte (Nutzungs-)Interessen an urheberrechtlich geschützten Gegenständen, die eigentlich die dem Urheber zugewiesene Verwertungen des Werks betreffen. Das gilt insbesondere für die Wissenschaft, schließlich ist für die Erstellung neuer Werke die Auseinandersetzung mit bereits existierenden Werken erforderlich. Diese Nutzungsinteressen werden ebenfalls durch das Urheberrecht geregelt. Das Instrument hierfür sind die **urheberrechtlichen Schranken**. Diese ermöglichen als **Erlaubnisnormen** bestimmte Handlungen, die eigentlich in die alleinige Kompetenz der Rechteinhaber fallen, z. B. die Vervielfältigung, § 16 UrhG, und die öffentliche Zugänglichmachung, § 19a UrhG. Der Begriff der Schranke soll dabei sinnbildlich veranschaulichen, dass sie die Rechte des Urhebers, d. h. seine Ausschließlichkeitsrechte, beschränken.

2.2 Gesetzesentwicklung

Die erste Erlaubnis zugunsten des TDM, § 60d UrhG a. F., privilegierte ausschließlich die nicht-kommerzielle wissenschaftliche Forschung und basierte auf den Spielräumen der damals geltenden Richtlinien, insbesondere die

Rn. 3; EuGH, Urt. v. 16. 7. 2009 – C-5/08, Infopaq = GRUR 2009, 1041, S. 1044; ob ein einheitlicher europäischer Werkbegriff existiert oder nur hinsichtlich einzelner Werkarten besteht, ist allerdings umstritten, mehr dazu von Schulze in Dreier/Schulze, UrhG, 6. Auflage 2018, § 2 Rn. 22; Leistner in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 3 ff.

¹⁷ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 11 Rn. 1 f.

¹⁸ Vogel in Schricker/Loewenheim, UrhG, 6. Auflage 2020, Vor § 87a Rn. 20.

InfoSoc-RL aus 2001.¹⁹ Seit dem 17.04.2019 ist eine verpflichtende TDM-Erlaubnis auch auf europäischer Ebene vorgesehen, Art. 3 und 4 der Richtlinie zum Urheberrecht im digitalen Binnenmarkt (DSM-RL). Art. 3 betrifft dabei die wissenschaftliche Forschung und Einrichtungen des Kulturerbes, Art. 4 erfasst auch kommerzielle und sonstige Zwecke. Leitgedanke war, die bestehenden Rechtsunsicherheiten mittels verpflichtender Ausnahmen oder Beschränkungen zu beseitigen²⁰ und die Wettbewerbsposition der EU zu stärken. Den Schaden für die Rechteinhaber schätzt der europäische Gesetzgeber in jedem Fall als minimal ein.²¹ Den Interessen der Rechteinhabern sei mit dem Erfordernis eines rechtmäßigen Zugangs zu den Werken Genüge getan.²²

Die Mitgliedsstaaten mussten die Vorgaben der DSM-RL bis zum 7.6.2021 in nationales Recht umsetzen.²³ Dazu wurde der § 60d UrhG basierend auf Art. 3 der DSM-RL aktualisiert und um eine **allgemeine TDM-Schranke** ohne die Bindung an die nicht-kommerzielle wissenschaftliche Forschung, **§ 44b UrhG**, ergänzt. § 44b UrhG definiert das TDM in Abs. 1 als *„die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen“*. § 44b UrhG. § 44b Abs. 2 UrhG erlaubt das Erstellen von Vervielfältigungen rechtmäßig zugänglicher digitaler oder digitalisierter Werke für das TDM ohne engere Zweckbindung. Die Vervielfältigungen sind nach Abs. 2 zu löschen, *„wenn sie für das TDM nicht mehr erforderlich sind“*. Gem. Abs. 3 können Rechteinhaber (maschinenlesbare) Vorbehalte anbringen, um das TDM an ihren Werken zu unterbinden. **§ 60d UrhG** erweitert die Erlaubnisse nach § 44b UrhG zugunsten von **Forschungsorganisationen** oder -instituten, **Kulturerbe-Einrichtungen** (Bibliotheken, öffentlich zugängliche Museen, Archive und Einrichtungen aus dem Film- und Tonerbe) und **einzelnen Forschenden** (Abs. 2 und 3). In Abs. 4 wird geregelt, inwieweit die Vervielfältigungen zu Zwecken gemeinsamer wissenschaftlichen Forschung Dritten öffentlich zugänglich gemacht werden dürfen. Abs. 5 erlaubt innerhalb bestimmter Grenzen die Aufbewahrung der Vervielfältigungen. Abs. 6 enthält Vorgaben dazu, wie Rechteinhaber das TDM an ihren Schutzgegenständen verhindern können.

¹⁹ Im Detail basiert § 60d UrhG a.F. auf Art. 5 Abs. 3 lit. a InfoSoc-RL sowie Art. 10 Abs. 1 lit. d Vermiet- und Verleih-RL und Art. 6 Abs. 2 lit. b, 9 lit. b Datenbank-RL.

²⁰ Erw. 10, 11, 15 DSM-RL.

²¹ Erw. 17 DSM-RL.

²² Erw. 14 DSM-RL.

²³ Erw. 10 DSM-RL.

Das praktizierte Text- und Data-Mining in den digitalen Geisteswissenschaften

Zwei Anwendungsfälle aus den digitalen Geisteswissenschaften sollen einen exemplarischen, in seinen Abläufen möglichst repräsentativen Workflow abbilden. Dieser Workflow bildet die Grundlage der rechtlichen Überlegungen.

3.1 Zusammenfassung des Forschungsprozesses

Den zeitaufwändigsten Abschnitt des Forschungsprozesses bildet das Sammeln und Aufbereiten der Forschungsdaten. In der Regel müssen analoge Forschungsdaten **digitalisiert** werden und der Text mittels **Optical Character Recognition (OCR)** in ein Reintextformat (.txt-Dateien) extrahiert werden. Anschließend wird der Text in den meisten Fällen in sog. **Tokens** segmentiert. Tokens sind die Wörter (definiert durch die sie begrenzenden Leer- und Satzzeichen) und Satzzeichen eines Textes. Diese Tokens oder Tokensequenzen sind dann Gegenstand umfangreicher **Annotationen**,²⁴ die sowohl manuell als auch automatisch erfolgen können. Denkbare Annotationen sind beispielsweise die Wortarten oder auch inhaltliche Merkmale von Texten. Es ist in der Praxis üblich, einen kleinen Anteil der Daten manuell zu annotieren, um so

²⁴ Unter Annotationen werden in diesem Gutachten sowohl die Annotationen an die einzelnen Textsegmente verstanden als auch solche, die sich auf größere Korpuszusammenhänge beziehen. Diesbezüglich wird auch der Begriff der Metadaten genutzt. Die Abgrenzung der Begriffe wird jedoch auch teilweise, insbesondere in den jeweiligen Fachdisziplinen, unterschiedlich vorgenommen.

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 9–12. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.c>. License: CC BY 4.0

ein Referenzset (den sog. Goldstandard) zu bilden. An diesem können für viele Analysekategorien Algorithmen trainiert werden, die den restlichen Anteil der Forschungsdaten automatisiert verarbeiten. Aus den vorverarbeiteten Forschungsdaten wird anschließend ein sog. Korpus gebildet. Das **Korpus** stellt also die **Zusammenstellung der vollständig aufbereiteten Forschungsdaten dar**, d. h. es ist eine tabellarische Anordnung der Segmente inklusive der Annotationen,²⁵ und ist **Gegenstand** der eigentlichen **TDM-Analyse**. Im Anschluss an die Analyse werden die Ergebnisse ausgewertet.

3.2 Status quo in den Digital Humanities

Empirische Studien legen den grundsätzlichen Schluss nahe, dass eine geringe Forschungsaktivität auf eine restriktive rechtliche Zulassung von TDM-Forschung zurückzuführen ist.²⁶ Doch auch seit Schaffung der Erlaubnisnorm zugunsten des nicht-kommerziellen, wissenschaftlichen TDMs (§ 60d UrhG a. F.) zum 1.3.2018 ist die TDM-Forschung an urheberrechtlich geschützten Werken noch keine gängige Praxis.²⁷ Insofern liegt nahe, dass eine restriktive Rechtslage, wie sie die Forschungsaktivitäten negativ beeinflusst, nicht nur eine generelle Erlaubnis, sondern auch die genaue Ausgestaltung der Erlaubnis negativ beeinflusst. Konkret betrifft das zum einen die Frage, ob der gesamte Prozess vollständig abgedeckt ist und zum anderen den Umgang mit den Korpora nach Abschluss der Forschungsarbeiten.²⁸ Zusammenfassend bedarf nicht nur irgendeines Rechtsrahmens, der das TDM erlaubt, sondern eines adäquaten Rechtsrahmens. Diese Annahme wird durch die Erfahrung der Projektbeteiligten aus verschiedensten Forschungsprojekten gestützt, gerade im Bereich der digitalen Literaturwissenschaft (als Unterkategorie der digitalen Geisteswissenschaften) wird urheberrechtlich geschützte Literatur nur selten mit digitalen Methoden beforscht, denn dazu ist oft die Digitalisierung der Texte notwendig, die einen erheblichen Arbeitsaufwand beinhaltet, der oftmals nur dann als angemessen empfunden wird, wenn die Digitalisate auch projektübergreifend genutzt werden können.²⁹ Außerdem sinken aufgrund eingeschränkter

²⁵ Dieses Verständnis liegt jedenfalls dem folgenden Gutachten zugrunde. In den betroffenen Fachdisziplinen wird das teilweise anders betrachtet.

²⁶ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 196 f.; Andresen et al., ZfdG im Erscheinen 2022; Handke/Guibault/Vallbb in Schmidt/Dobrevá (Hrsg.), *New avenues for electronic publishing in the age of infinite collections and citizen science*, 127 ff; Haucap/Loebert/Spindler/Thorwarth, *Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht*, S. 45 f; ebenso Raue, GRUR 2017, S. 11, 13.

²⁷ Z. B. in den Projekten des Trier Center for Digital Humanities, so Raue/Schöch, RuZ 2020, S. 118, 118 f.

²⁸ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 196 f.

²⁹ Ebd.

Publikationsmöglichkeiten und fehlender Open-Access-Konformität die Förderchancen für die entsprechenden Projekte.³⁰ Aktuell beschränken sich die digitalen Geisteswissenschaften auf gemeinfreie Literatur, die ca. zwischen 1800 und 1920 entstand, was darauf zurückgeführt wird, dass ab dieser Grenze viele Werke gemeinfrei sind und ältere Literatur (mit zumutbarem Aufwand) nur schwer mittels OCR zu digitalisieren ist.³¹ Die Nutzung gemeinfreier Werke birgt also den Vorteil, dass die erstellten Korpora projektübergreifend genutzt und digital geteilt werden dürfen.³²

3.3 Anwendungsfälle

3.3.1 Erster Anwendungsfall: Unzuverlässiges Erzählen

Der erste Anwendungsfall erforscht das Phänomen des unzuverlässigen Erzählens der Narratologie. Von unzuverlässigem Erzählen wird gesprochen, wenn dem Erzähler eines literarischen Textes nicht getraut werden kann, z. B. aufgrund von Fehlinformation, Fehlinterpretation oder allgemeiner Unglaubwürdigkeit.³³ Im vorliegenden Anwendungsfall wird gerade die faktenbezogene Unzuverlässigkeit untersucht. Die Untersuchung erfolgt mittels automatischer und manueller Annotationen von Primär- und Sekundärtexten, die beantworten sollen, welche textuellen Indikatoren und textexternen Informationen bei der Feststellung unzuverlässigen Erzählens auf welche Weise zusammenspielen. Außerdem wird die Plausibilität unterschiedlicher Interpretationen unzuverlässig erzählter Texte verglichen.³⁴

3.3.2 Zweiter Anwendungsfall: Forschungsliteratur und Wissenschaftssprache

Der zweite Anwendungsfall widmet sich der Wissenschaftssprache in den Geisteswissenschaften am Beispiel der Fächer Literaturwissenschaft, Linguistik und Philosophie. Er baut darauf auf, dass zahlreiche Unterschiede zwischen den Wissenschaftssprachen von Literaturwissenschaft und Linguistik existieren,³⁵ z. B. werden in linguistischen Texten (unter anderem) mehr Textkommentare (*in der vorliegenden Arbeit ...*) verwendet und Argumentationsschritte

³⁰ Ebd.

³¹ Schöch et al., ZfdG 2020 Rn. 1; Döhl, RuZ 2020, S. 195, 207; Raue/Schöch, RuZ 2020, S. 118, 118 f; Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 196 f.

³² Ebd.

³³ Jacke, Systematik unzuverlässigen Erzählens, S. 1 f; Shen in Hühn/Meister/Pier/Schmid (Hrsg.), *the living handbook of narratology*.

³⁴ Vgl. dazu ausführlich Andresen et al., ZfdG im Erscheinen 2022.

³⁵ Andresen, Datengeleitete Sprachbeschreibung mit syntaktischen Annotationen.

expliziter versprachlicht (*zu dem Schluss kommen, dass ...*). Als Datengrundlage dient eine Auswahl von wissenschaftlichen Zeitschriftenartikeln. Gerade weil diese Textsorte in den Fächern breit rezipiert wird und strengen Qualitätskriterien unterliegt, stellen sie eine geeignete Stichprobe für die Wissenschaftssprache dieser Fächer dar.³⁶

³⁶ Vgl. auch hierzu ausführlich Andresen et al., ZfdG im Erscheinen 2022.

KAPITEL 4

Rechtliche Beurteilung

Im Folgenden wird der Rechtsrahmen daraufhin untersucht, ob er den bereits in Kürze zusammengefassten Forschungsprozess der beiden Anwendungsfälle (vollständig) abdeckt.

4.1 Berechtigte

Auf § 60d UrhG können sich, allgemein gesagt, Personenkreise berufen, die nicht-kommerzielle wissenschaftliche Forschung betreiben. Forschungsorganisationen können sich dabei auf § 60d Abs. 2 S. 1, 2 UrhG stützen, Kulturerbe-Einrichtungen wie Bibliotheken, öffentlich zugängliche Museen, Archive und Einrichtungen aus dem Film- oder Tonerbe auf § 60d Abs. 3 Nr. 1 UrhG, Einzelforschende auf § 60d Abs. 3 Nr. 2 UrhG.

Akteure der Wirtschaft sowie andere Personenkreise wie z. B. die Presse wurden erst nach Umsetzung der DSM-RL in das nationale Urheberrecht zum 7.6.2021 einbezogen, § 44b UrhG. Dazu zählen auch Zusammenarbeiten von öffentlichen und privaten Trägern (Public-Private-Partnerships), sofern den privaten Trägern ein Einfluss auf die Forschungsorganisation sowie ein bevorzugter Zugriff auf die Forschungsergebnisse zukommt (vgl. § 60d Abs. 2 S. 3 UrhG).

4.2 Schutz des Ursprungsmaterials

Die Quellen, die analysiert werden sollen, können als Werke gem. § 2 Abs. 2 UrhG urheberrechtlich geschützt sein, wenn sie die notwendige

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 13–40. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.d>. License: CC BY 4.0

Schöpfungshöhe aufweisen.³⁷ Texte können u. a. als Sprach- bzw. Schriftwerke geschützt sein, § 2 Abs. 1 Nr. 1, Abs. 2 UrhG. Daneben ist bei Sprachwerken die „*kleine Münze*“ anerkannt,³⁸ es sind also geringere Anforderungen an die Schöpfungshöhe zu stellen.

Im ersten Anwendungsfall werden vorwiegend Romane, aber auch Kurzgeschichten, untersucht. Solche klassischen literarischen Werke sind regelmäßig urheberrechtlich geschützt.³⁹ Eine Besonderheit gilt im ersten Anwendungsfall für die Kurzgeschichte „*Rosalie geht sterben*“ in *Daniel Kehlmanns* Sammlung „*Ruhm*“ (2009 erschienen). Die in der Sammlung enthaltenen Kurzgeschichten sind inhaltlich miteinander verwoben, sodass nahe liegt, dass sie erstens jeweils einzeln als Sprachwerke nach §§ 2 Abs. 1 Nr. 1, Abs. 2 UrhG, sowie zweitens auch insgesamt als Sammelwerk i. S. d. § 4 Abs. 1 UrhG geschützt sind. Insofern sind bei denselben Quellen u. U. mehrere Werke Bezugsobjekt. Aber auch Forschungsliteratur, wie sie im zweiten Anwendungsfall untersucht wird, kann urheberrechtlich geschützt sein. Die wissenschaftlichen Erkenntnisse als solche sind allerdings frei.⁴⁰ Geschützt sein kann im Bereich von Forschungsliteratur insofern nur die Darstellungsweise,⁴¹ nicht aber der vermittelte Inhalt.⁴² Daneben fehlt gerade der sog. „*gewöhnlichen wissenschaftlichen Ausdrucksweise*“

³⁷ Wiebe in Spindler/Schuster, Recht der elektronischen Medien, § 2 UrhG Rn. 2.

³⁸ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 85; OLG Nürnberg, Urt. v. 27. 3. 2001 – 3 U 3760/00, Dienstanweisung = GRUR-RR 2001, 225, S. 226 f.

³⁹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 86.

⁴⁰ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 41, 93; BGH, Urt. v. 21.11.1980 – I ZR 106/78, Staatsexamensarbeit = GRUR 1981, 352, S. 353; BGH, Urt. v. 27.3.1963 – I b ZR 129/6, Rechenschieber = GRUR 1963, 633, S. 634. Wissen wird deswegen als öffentliches Gut bezeichnet, das dem Urheberrecht nicht zugewiesen ist, so u. A. Moltke, Das Urheberrecht an den Werken der Wissenschaft, S. 57.

⁴¹ Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 229.; BGH – Urt. v. 21.11.1980 – I ZR 106/78, Staatsexamensarbeit = GRUR 1981, 352 ff., S. 353.; BGH – Urt. v. 17.04.1986 – I ZR 213/83, Anwaltsschriftsatz = GRUR 1986, 739 ff., S. 740.

⁴² Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 13; Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn, 47; Bullinger in Wandtke/Bullinger, UrhG, § 2 Rn, 19; Geiger/Frosio/Bulanyenko, IIC 2018, S. 814, 817; EuGH, ECLI:EU:C:2018:899, Levola Hengelo/Smilde Foods = GRUR 2019, 73, S. 74 Rn. 40. Einzige Ausnahme dieses Grundsatzes können fiktionale Geschichten sein, die – sofern sie schon eine ausreichende Gestalt angenommen haben – auch ihres Inhalts nach geschützt sein können, vgl. dazu Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 75, 78.

oftmals die notwendige Individualität.⁴³ Insofern ist der Spielraum für die Erreichung dieser Individualität häufig gering, dennoch muss in der Praxis im Zweifel von einer Schutzfähigkeit ausgegangen werden.

Gerade bei Sprachwerken ist zu beachten, dass das Werk gerade durch das Mittel der Sprache zum Ausdruck kommt,⁴⁴ geschützt ist zwar die Darstellungsweise bzw. die Argumentationsstruktur, bei fiktionalen Geschichten auch teilweise inhaltliche Elemente,⁴⁵ sofern sie im Werk manifestiert sind, nicht geschützt ist hingegen die genaue Formgestaltung des Textes bzw. Layout.

4.3 Zugriff auf das Ursprungsmaterial: Rechtmäßiger Zugriff zwischen Verträgen und technischen Schutzmaßnahmen

Voraussetzung der §§ 60d, 44b UrhG sowie Art. 3, 4 DSM-RL ist ausdrücklich, dass die beforschten Schutzgegenstände für die Nutzenden **rechtmäßig zugänglich sind**.⁴⁶ Dieser rechtmäßige Zugang liegt etwa vor, wenn das Werk als körperliches Werkexemplar, z. B. in analoger Buchform, vorliegt.⁴⁷ Dasselbe gilt für online frei verfügbare Daten oder Werke, zu denen der Zugang über Lizenz erlangt wurde.⁴⁸ Nutzungsbedingungen sogenannter Digital-Rights-Management Systeme, über die z. B. die Bereitstellung von E-Books erfolgen kann, beinhalten allerdings gelegentlich Beschränkungen auf eine reine Online-Nutzung und ermöglichen nur den Download kleiner Anteile. Derartige Beschränkungen sind nicht (mehr) rechtlich durchsetzbar, wie im Folgenden dargestellt wird.

4.3.1 Vertragsfestigkeit der TDM-Schranke

§ 60g Abs. 1 UrhG besagt, dass sich **Rechteinhaber nicht auf Vereinbarungen, die zum Nachteil der Nutzungsberechtigten u. a. von § 60d UrhG**

⁴³ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 93; Kuschel, OdW 2020, S. 43, 44; BGH, Urt. v. 21.11.1980 – I ZR 106/78, Staatsexamensarbeit = GRUR 1981, 352, S. 355; BGH, Urt. v. 29.03.1984 – I ZR 32/82, Ausschreibungsunterlagen = GRUR 1984, 659, S. 663.

⁴⁴ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 81.

⁴⁵ Zu inhaltlichen Elementen einer Fabel Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 75, 103; BGH, Urt. v. 29.04.1999 – I ZR 65/96, Laras Tochter = GRUR 1999, 984, S. 987.

⁴⁶ Das galt auch bereits vor Umsetzung der DSM-RL, BT-Drs. 18/12329, S. 41; DSM-RL Erw. 18 UAbs. 2; BT-Drs. 19/27426, S. 88, 96.

⁴⁷ Raue, ZUM 2021, S. 793, 800.

⁴⁸ DSM-RL Erw. 14; BT-Drs. 18/12329, S. 41; BT-Drs. 19/27426, S. 88, 96; Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 44b Rn. 8, § 60d Rn. 2; Raue, ZUM 2021, S. 793, 800.

abweichen, berufen können. TDM-Nutzungen durch die nicht-kommerzielle wissenschaftlichen Forschung gem. § 60d UrhG sind also vertragsfest ausgestaltet. Etwas anderes gilt für TDM außerhalb der privilegierten Zwecke der nicht-kommerziellen wissenschaftlichen Forschung, § 44b Abs. 3 UrhG. In kommerziellen Kontext oder zu anderen, außerhalb der in § 60d privilegierten Zwecke, dürfen Rechteinhaber Vorbehalte anbringen, die das TDM untersagen. Die Vorbehalte müssen in angemessener Art und Weise ausgesprochen werden, sind die Werke online zugänglich, ist dazu eine Maschinenlesbarkeit Maßstab, § 44b Abs. 3 UrhG.⁴⁹ Es können also einseitige Vorbehalte ausgesprochen werden und auch vertragliche Regelungen getroffen werden, z. B. zu einer Vergütung.

4.3.2 Durchsetzbarkeit der urheberrechtlichen Schranke gegenüber technischen Schutzmaßnahmen, §§ 60g, 95b UrhG

Technische Systeme, die den vollständigen Download technisch verhindern, sind zwar zulässig, aber nicht rechtlich durchsetzbar. Konkret besteht ein Anspruch auf Zurverfügungstellung eines geeigneten Mittels zur Umgehung der technischen Barriere.⁵⁰ Dies gilt seit dem 7.6.2021 auch für digitale Zugänglichmachungen auf Vertragsbasis im Rahmen des TDM, § 95b Abs. 3 Nr. 1, 5 UrhG, jedenfalls dann, wenn Forschungs- und Kulturerbe-Einrichtungen oder die allgemeinen Personenkreise (§ 44b UrhG) betroffen sind.⁵¹ Es ergibt sich daraus allerdings kein „*right to hack*“.⁵²

4.3.3 Andere Umgehungsmöglichkeiten

Rechteinhaber können ihre Schutzgegenstände in Kontexten der nicht-kommerziellen wissenschaftlichen Forschung ausschließlich durch

⁴⁹ BT-Drs. 19/27426, S. 89.

⁵⁰ Götting in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 95b Rn. 21.

⁵¹ Raue, ZUM 2021, S. 793, 801 f. Das wurde von Art. 7 Abs. 2 DSM-RL vorgegeben. Etwas anderes galt auf Grundlage des § 95b Abs. 3 UrhG a. F. Vgl. dazu ausführlich Specht-Riemenschneider, Diktat der Technik, S. 396; Götting in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 95b Rn. 26; Specht-Riemenschneider in Dreier/Schulze, UrhG, 6. Auflage 2018, § 95b Rn. 18; Wandtke/Ohst in Wandtke/Bullinger, UrhG, § 95b Rn. 49; Czychowski in Fromm/Nordemann, UrhG, § 95b Rn. 26 f.; Dreier, ZUM 2002, S. 28, 37; Spindler, GRUR 2002, S. 105, 117, 119. In Bezug auf das vormalig geltende EU-Recht sowie einen diesbezüglich defizitären RL-Entwurf zurecht kritisch Ducato/Stowe, IIC 2019, S. 649, 666 Schlechter gestellt sind dementsprechend Individualforschende (citizen science).

⁵² Wandtke/Ohst in Wandtke/Bullinger, UrhG, § 95b Rn. 16.

Maßnahmen zur Sicherheit und Integrität ihrer Netze oder Datenbanken sichern, § 60d Abs. 6 UrhG. Relevant ist das vor allem im Online-Bereich, wenn Webseiten bei einer systematischen und automatischen Abfrage überlastet werden. Die Schutzmaßnahmen dürfen das TDM aber nicht allgemein unterbinden, sondern müssen sich auf das zum Erhalt der Funktionsfähigkeit notwendigen Maß beschränken.⁵³

Kurz zusammengefasst...

- Um Werke mit TDM-Methoden in urheberrechtlich relevanter Weise nutzen zu dürfen, muss zunächst ein rechtmäßiger Zugang erworben werden.
- Vertraglich kann von § 60d UrhG nicht zulasten des Schrankenbegünstigten abgewichen werden.
- Technische Schutzmaßnahmen dürfen die Erlaubnisse nach § 60d UrhG durch Forschungs- und Kulturerbe-Einrichtungen sowie nach § 44b UrhG nicht beschränken.
- Nutzungen, die der § 60d UrhG erlaubt, können nur insoweit verhindert werden, wie es die Funktionsfähigkeit der Systeme erfordert, § 60d Abs. 6 UrhG.

4.4 Digitalisierung analoger Quellen

Analoge Forschungsquellen müssen zunächst digitalisiert werden. Im Folgenden wird dargelegt, welche Rechte bei einer Digitalisierung verletzt werden und ob diese Verletzungshandlungen durch die TDM-Schranken gerechtfertigt sind.

4.4.1 Korrekte Digitalisierung

In Bezug auf die korrekte Digitalisierung analoger Quelle kommt in Betracht, dass sie vervielfältigt, § 16 UrhG und/oder bearbeitet werden, § 23 UrhG.

4.4.1.1 Vervielfältigungen, § 16 Abs. 1 UrhG

Unter einer Vervielfältigung ist „jede körperliche Festlegung eines Werks“ zu verstehen, „die geeignet ist, das Werk den menschlichen Sinnen auf irgendeine Weise

⁵³ DSM-RL Erw. 16; BT-Drs. 19/27426, S. 97 f.

unmittelbar oder mittelbar zugänglich zu machen“.⁵⁴ Bei Digitalisierungen handelt es sich in aller Regel um Vervielfältigungen i. S. d. § 16 Abs. 1 UrhG,⁵⁵ denn das Werk wird körperlich auf ein anderes Medium übertragen und ist für den Menschen dann mit digital wahrnehmbar. Das Vervielfältigungsrecht ist ein Verbotsrecht, d. h. es steht ohne gesonderte Erlaubnisnorm oder vertragliche Gestattung nur dem Rechteinhaber zu, sein Werk zu vervielfältigen.⁵⁶

Zum Zwecke der Digitalisierung urheberrechtlich geschützter Quellen ist also erforderlich, dass eine gesetzliche Schranke eingreift. Es stellt sich deswegen die Frage, ob die TDM-Schranken des §§ 60d, 44b Abs. 1, 2 S. 1 UrhG auch die (erstmaligen) Digitalisierung (analoger Werke) erfassen oder ob nur originär digital zugängliche Werke untersucht werden dürfen. Art. 2 Nr. 2 DSM-RL definiert das TDM als „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“ definiert ist. Das impliziert, dass die DSM-RL davon ausgeht, dass die Ursprungsmaterialien bereits in digitaler Form vorliegen müssen. Die DSM-Richtlinie geht auch in Erw. 8 davon aus, dass das Ursprungsmaterial bereits in digitaler Form vorliegt. Dennoch erlaubt die DSM-RL die Verwertungshandlungen „zum Zwecke des TDM“, Art. 3 Abs. 1, Art. 4 Abs. 1 DSM-RL, was wiederum ein weites Verständnis zugrunde legt, d. h. erst die Analyse erfolgt an digitalen Inhalten, die zuvor rechtmäßigerweise digitalisiert werden können. In Bezug auf §§ 44b, 60d UrhG n. F. stellt die Gesetzesbegründung der deutschen Bundesregierung allerdings klar, dass die Digitalisierung analoger Quellen erfasst ist.⁵⁷

4.4.1.2 Bearbeitungen, § 23 UrhG

Unter Bearbeitungen werden Eingriffe in das Werk verstanden, die dem Werk dienen und es einem veränderten Zweck anpassen sollen, indem es z.

⁵⁴ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 16 Rn. 5; BGH, Urt. v. 18.05.1955 – I ZR 8/54, Magnettonbänder II = GRUR 1955, 492, S. 494.

⁵⁵ Loewenheim in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 16 Rn. 20; Heerma in Wandtke/Bullinger, UrhG, § 16 Rn. 16; BGH, Urt. v. 16.4.2015 – I ZR 69/11, Elektronische Leseplätze II = GRUR 2015, 1101, S. 821 Rn. 25.

⁵⁶ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 16 Rn. 18.

⁵⁷ BT-Drs. 19/27426, S. 88. Wenn man davon ausgeht, dass die DSM-RL die Digitalisierung nicht erfasst, wäre die Erlaubnis der erstmaligen Digitalisierung im nationalen Recht jedenfalls durch eine überschießende Umsetzung der DSM-RL auf Grundlage der InfoSoc-RL geregelt. Das ermöglichen Art. 25 DSM-RL und Art. 5 Abs. 3 lit. a InfoSoc-RL im Rahmen des § 60d UrhG.

B. übersetzt, verfilmt oder auf einen anderen Werkstoff übertragen wird.⁵⁸ Umgestaltungen sind hingegen Nutzungen des Werks in abgeänderter Form, die dem Originalwerk nicht dienen.⁵⁹ Die Differenzierung hat jedoch keine weiteren Auswirkungen, weswegen im Folgenden darauf verzichtet wird.

Auch Bearbeitungen i. S. d. § 23 Abs. 1 UrhG müssen die persönliche geistige Schöpfung betreffen.⁶⁰ Dabei ist nicht jede Änderung auch ausschlaggebend, d. h. es kann auch abgewandelte Vervielfältigungen geben, die noch nicht als Bearbeitung einzuordnen sind.⁶¹

Bei Digitalisierungen werden die schutzfähigen Elemente der Werke, die gerade in der Sprache, nicht aber in einem (Datei-)Format liegen,⁶² nicht verändert.⁶³ Dass in solchen Fällen keine Bearbeitung gegeben ist, stellt auch § 23 Abs. 3 UrhG klar. Ihm zufolge finden § 23 Abs. 1, 2 UrhG, d. h. die Regelungen zu Bearbeitung urheberrechtlicher Werke, keine Anwendung, wenn es sich um „*ausschließlich technisch bedingte Änderungen*“ handelt. Gemeint sind wohl in erster Linie Formatänderungen.⁶⁴ Genau diese Formatänderungen werden bei Digitalisierungen vorgenommen.

4.4.2 Fehlerhafte Digitalisierungen

Bei der Digitalisierung können diverse Fehler wie Verzerrungen, Schatten (oder auch mitgescannte Finger) erfolgen. Eine Bearbeitung (§ 23 Abs. 1 UrhG) ist auch dann gegeben, wenn sich der vom Urheber bezweckte geistig-ästhetische Gesamteindruck verändert.⁶⁵ Bei Sprachwerken liegt die Schöpfung etwa in

⁵⁸ Loewenheim in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 23 Rn. 3, 4; Wiebe in Spindler/Schuster, Recht der elektronischen Medien, § 23 UrhG Rn. 4; Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 23 Rn. 22; BGH, Urt. v. 08.11.1989 – I ZR 14/88, Bibelreproduktion = GRUR 1990, 559, S. 673.

⁵⁹ A. A. Bullinger in Wandtke/Bullinger, UrhG, § 23 Rn. 3.

⁶⁰ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 23 Rn. 24; BGH, Urt. v. 7. 2. 2002 – I ZR 304/99, Unikatrahmen = GRUR 2002, 532, S. 534.

⁶¹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 16 Rn. 10; BGH, Urt. v. 10.12.1987 – I ZR 198/85, Vorentwurf II = GRUR 1988, 533, S. 535; BGH, Urt. v. 28.02.1991 – I ZR 88/89, Explosionszeichnungen = GRUR 1991, 529, S. 530.

⁶² Vgl. dazu bereits o. S. 9 f.

⁶³ Loewenheim in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 23 Rn. 6 mit weiteren Nachweisen; vgl. auch die ständige Rechtsprechung BGH, Urt. v. 21.11.1952 – I ZR 56/52, Magnettonbänder I = GRUR 1953, 140, S. 141 f.

⁶⁴ BT-Drs. 18/12329, S. 31; Hegemann/Naldony in Hoeren/Sieber/Holznel, Multimediarecht Hdb, § 60d Rn. 138.

⁶⁵ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 23 Rn. 24.

der Gedankenführung des Inhalts und in der Art der Sammlung des Stoffs.⁶⁶ Sofern eine Verzerrung o. Ä. nur einzelne Wörter oder Abschnitte betrifft, ist der geistig-ästhetische Gesamteindruck jedenfalls zumeist nicht verändert, was zur Folge hat, das darin nur eine Vervielfältigung i. S. d. § 16 Abs. 1 UrhG zu sehen ist. Möglich erscheint das Vorliegen einer Bearbeitung gegebenenfalls, wenn Textabschnitte, die relevant für das Gesamtverständnis sind, unlesbar sind. Im deutschen UrhG gilt allerdings der Grundsatz der Herstellungsfreiheit von Bearbeitungen.⁶⁷ Dass ein Werk bearbeitet wurde, wird insofern ohnehin erst dann relevant, wenn das bearbeitete Werkexemplar weiter verwertet werden soll.⁶⁸

Grundsätzlich ist bei diesen Fehlern auch an eine Entstellung nach § 14 UrhG zu denken. Im Bereich der gesetzlichen Schrankenbestimmungen wird diese Norm durch das sog. Änderungsverbot gem. § 62 UrhG verdrängt,⁶⁹ der bestimmte Änderungen im Rahmen urheberrechtlicher Schrankenbestimmungen erlaubt.⁷⁰ Im Falle der Herstellung bearbeiteter Werkstücke wird § 62 UrhG allerdings durch die Herstellungsfreiheit der Bearbeitung beschränkt.⁷¹ Das bedeutet, dass auch die Herstellung eines veränderten Werkstücks solange möglich ist, wie es eine Bearbeitung darstellt. In Ergänzung an die bisherigen Ausführungen ist insofern auch eine Änderung zulässig.

Kurz zusammengefasst...

Die Digitalisierungsvorgänge sind vom geltenden Urheberrecht abgedeckt, unabhängig davon, ob sie korrekt oder fehlerbehaftet sind.

⁶⁶ Loewenheim/Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 2 Rn. 78.; BGH, Urt. v. 11. 4. 2002 – I ZR 231/99, Technische Lieferbedingungen = GRUR 2002, 958, S. 959; BGH, Urt. v. 21.11.1980 – I ZR 106/78, Staatsexamensarbeit = GRUR 1981, 352, S. 353; BGH, Urt. v. 07.12.1979 I ZR 157/77, MonumentaGermamae Historica = GRUR 1980, 227, S. 930.

⁶⁷ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 23 Rn. 33; Loewenheim in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 23 Rn. 18. Eine Ausnahme bildet z. B. das Datenbankwerk, § 23 Abs. 2 UrhG.

⁶⁸ Dazu kommt es im späteren Forschungsablauf, S. dazu S. 18.

⁶⁹ Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 62 Rn. 1.

⁷⁰ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 62 Rn. 1; Peukert in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 62 Rn. 9 ff.

⁷¹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 62 Rn. 8; Bullinger in Wandtke/Bullinger, UrhG, § 62 Rn. 4.

4.5 Texterkennung mithilfe von OCR

Ein wichtiger Vorverarbeitungsschritt, der sich an die Digitalisierung analoger Quellen anschließt, ist die Extrahierung des Reintextes aus den Scans mittels OCR, die zu diesem Zeitpunkt nur Bilddateien darstellen. Ziel ist, das Ausgangsmaterial maschinenlesbar zu machen, sodass es später automatisiert analysiert werden kann. Im Grunde werden bei der OCR-Analyse Pixel-Anhäufungen als Buchstaben, Zahlen oder andere Zeichen erkannt. Zu diesem Zwecke werden die PDF- oder sonstigen Dokumente mithilfe geeigneter Software untersucht, woraufhin zunächst OCR-Rohdaten entstehen. Diese beinhalten wiederum Informationen über die erkannten Zeichen und deren geschätzte Position im Eingabebild, die Sicherheit, mit der das Tool die Entscheidung getroffen hat, sowie den erkannten Text selbst. Dieser wird dann zumeist zusätzlich als Reintext, z. B. in eine txt-Datei, extrahiert.

4.5.1 Korrekte Texterkennung

Sofern die OCR den Text korrekt erkennt, handelt es sich um eine Vervielfältigung, § 16 Abs. 1 UrhG. Gerade, wenn der Text in ein Reintextformat extrahiert wird, wird das Werk zwar nicht in dem Layout dargestellt, wie er im originalen Werk bzw. dem Scan zu erkennen war. Dennoch wird das Werk in seinem geistig-schöpferischen Gehalt, der zumeist nicht in der Aufteilung des Textes auf einer Seite bzw. der Formatierung liegt,⁷² durch OCR nicht verändert.⁷³ Die Veränderungen gehen insofern nicht darüber hinaus, was auch bei einer einfachen Digitalisierung wie etwa einem Scan geschieht, die anerkanntermaßen nur eine Vervielfältigung ist.⁷⁴ Das gilt erst recht, weil nicht nur identische, sondern auch leicht abgewandelte Kopien Vervielfältigungen i. S. d. § 16 Abs. 1 UrhG darstellen, zumal der Schutzgegenstand des Werks nicht in der optischen Darstellung liegt, sondern in der Sprache.⁷⁵ Auch dies spricht für eine gleiche Bewertung wie die Digitalisierung.⁷⁶ Das gleiche ergibt § 23 Abs. 3 UrhG, der klarstellt, dass Formatänderungen als (ggf. leicht abgewandelte) Vervielfältigungen gelten. Insofern ist die Textextrahierung unter den Voraussetzungen der §§ 44b, 60d UrhG erlaubt.

⁷² Vgl. dazu bereits S. 9 f.

⁷³ So auch Leventer, Google Book Search und vergleichendes Urheberrecht, S. 232.

⁷⁴ Vgl. o. Fn. 55.

⁷⁵ Vgl. o. Fn. 61. Dazu zählt explizit auch das Abschreiben eines Textes, Loewenheim in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 16 Rn. 9.

⁷⁶ Ebenso Specht, OdW 2018, S. 285.

4.5.2 Fehlerhafte Texterkennungen

Fehlerhafte Texterkennungen können zum einen Bearbeitungen gem. § 23 Abs. 1 UrhG sein, denn je nachdem, wie weitreichend die Fehler in der Texterkennung sind, ist denkbar, dass der geistige Gehalt eines Werks oder Werkteils verändert wird.⁷⁷ Das kommt etwa dann in Betracht, wenn Wörter falsch erkannt werden und sich auf diese Weise der Sinn von Texten ändert. Minimale Fehlerkennungen genügen hierfür nur in den seltensten Fällen. In den Fällen, in denen die fehlerhaft erkannten Abschnitte den geistigen Gehalt und damit das urheberrechtliche Schutzgut berühren, handelt es sich um Bearbeitungen, § 23 Abs. 1 UrhG, bei unwesentlichen Änderungen liegt nur eine abgewandelte Vervielfältigung gem. § 16 UrhG vor, die unter den Voraussetzungen der §§ 44b, 60d UrhG erlaubt ist.

Fehlerhafte Texterkennungen können zum anderen auch als Änderungen nach § 62 UrhG einzuordnen sein. Im Anwendungsbereich der Bearbeitung gem. § 23 UrhG wird das Änderungsverbot allerdings aufgrund der Herstellungsfreiheit von Bearbeitungen verdrängt.⁷⁸

4.5.3 Händische Übertragungen

In manchen Fällen liefert die OCR-Software keine guten Ergebnisse, sodass händisch nachgebessert wird. Bei der rechtlichen Bewertung dieser händischen Nachbesserungen muss je nach Umfang der Nachbesserungen differenziert werden. Bei einzelnen Korrekturen falsch extrahierter Passagen wird nicht das eigentliche Werk verändert, sondern der fehlerhafte Output der OCR-Software. Denkbar ist allerdings, dass die händischen Korrekturen in manchen Fällen Vervielfältigungen i. S. d. § 16 Abs. 1 UrhG darstellen. Das ist dann der Fall, wenn ein urheberrechtlich geschützter Werkteil vervielfältigt wird.⁷⁹ Bei einzelnen Wörtern oder Silben ist das abzulehnen, während bereits kürzere Wortfolgen dem Urheberrechtsschutz unterliegen können.⁸⁰ Dabei gilt, dass die Wahrscheinlichkeit eines Urheberrechtsschutzes mit der Länge einer Wortfolge steigt.⁸¹ Je nachdem, in welchem Umfang die Korrekturen erfolgen, z. B. beim Abtippen ganzer Sätze oder Seiten, können Vervielfältigungen nach § 16 S. 1 UrhG nicht gänzlich ausgeschlossen werden.⁸²

⁷⁷ So auch Leventer, Google Book Search und vergleichendes Urheberrecht, S. 232.

⁷⁸ Vgl. dazu bereits auf S. 13 f.

⁷⁹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 16 Rn. 9.

⁸⁰ EuGH, Urt. v. 16. 7. 2009 – C-5/08, Infopaq = GRUR 2009, 1041.

⁸¹ Jotzo, RuZ 2020, S. 128, 133.

⁸² Gleiches gilt, wenn ein Text einfach händisch abgeschrieben wird, weil es damit körperlich festgelegt wird, vgl. Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 16 Rn. 7.

Bei der sich an diese Überlegung anschließenden Frage danach, ob händische Nachbesserungen auch von § 60d UrhG erlaubt werden, kommt es darauf an, ob § 60d UrhG neben automatischen Prozessen auch einzelne manuelle Vorgänge erlaubt. Seinem Wortlaut nach erfasste jedenfalls § 60d Abs. 1 Nr. 1 UrhG a. F. Vorgänge, die *auch* automatisiert ablaufen, um das Material später zu TDM-Zwecken automatisch auszuwerten, demnach war jedenfalls die Vervielfältigung nach altem Recht nicht ausschließlich auf automatisierte Vorgänge beschränkt.⁸³ Die Legaldefinition in Art. 2 Abs. 1 Nr. 2 DSM-RL verwendet die Begrifflichkeit der *automatisierten* Daten- bzw. Textanalyse. Das bedeutet freilich nicht, dass auch alle Einzelhandlungen zur Vorbereitung automatisiert erfolgen müssen. Dasselbe gilt für die neue Fassung des § 60d UrhG, die sich zwar auf automatisierte Vorgänge bezieht, die Vervielfältigungen selbst sind allerdings nach dem Wortlaut nicht zwingend automatisiert. Dies spricht dafür, jedenfalls einzelne händische Vervielfältigungen zur Vorbereitung des Forschungsmaterials zu erfassen. Sofern man eine Auslegung des deutschen Rechts über die genauen Inhalte der DSM-RL vornimmt, können diese auf Art. 5 Abs. 1 lit. a InfoSoc-RL gestützt werden, Art. 25 DSM-RL.

Kurz zusammengefasst...

Bei der OCR-Digitalisierung erfolgen urheberrechtliche Vervielfältigungen. Diese sind im Rahmen der §§ 60d, 44b Abs. 1 UrhG erlaubt. Händische Korrekturen sind in der Regel nur dann, wenn sie nicht nur einzelne Wörter betreffen, als Vervielfältigungen, § 16 UrhG, zu bezeichnen, die auch von §§ 60d, 44b Abs. 1 UrhG erlaubt werden.

4.6 Segmentierung

Bei der Segmentierung werden die Eingabedaten in Tokens zerlegt (Abbildung 1). Dabei wird der Reintext zunächst dupliziert, anschließend in die Tokens aufgespalten und in dieser neuen Repräsentationsform gespeichert. Auch diese Verarbeitung, die auch die Textstruktur betrifft, wirft Fragen nach der urheberrechtlichen Zulässigkeit auf.

4.6.1 Vervielfältigungen und Bearbeitungen, §§ 16 Abs. 1, 23 Abs. 1 UrhG

Weil jedenfalls das äußere Erscheinungsbild eines Textes in erheblicherem Maß verändert wird muss das Vorliegen einer Bearbeitung nach § 23 Abs. 1 UrhG

⁸³ Vgl. Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, § 60d Rn. 7; Specht, OdW 2018, S. 285, 286.

Eingabetext	Tokenisiert (ein Token pro Zeile, Satztrennung durch Leerzeile)
Wir fragen die Arbeitsleute, ob sie uns nicht die Spur eines Tigers anzugeben wüßten.	Wir fragen die Arbeitsleute , ob sie uns nicht die Spur eines Tigers anzugeben wüßten .

Abbildung 1: Segmentierung. Quelle: Eigene Darstellung, Beispielsatz aus „Eine Frauenfahrt um die Welt“ von Ida Pfeiffer (1850), CC BY 4.0 International.

in Erwägung gezogen werden. Durch die Kopie einzelner kurzer Textsegmente ist das urheberrechtliche Werk zumeist noch nicht betroffen, etwas anderes gilt jedoch durch die Übernahme aller Textsegmente, die wiederum das gesamte Werk oder jedenfalls schutzfähige Textabschnitte betrifft.

Ob das urheberrechtliche Werk tatsächlich bearbeitet wird, ist zwar nur im Einzelfall zu entscheiden. Bei textbasierten Werken ist es aber eher unwahrscheinlich: Gerade literarische Werke zeichnen sich in aller Regel⁸⁴ dadurch aus, dass der geistige Inhalt durch das Mittel der Sprache zum Ausdruck kommt.⁸⁵ Diese Sprache wird bei der Segmentierung von Fließtexten⁸⁶ nicht geändert, denn auch als segmentierte Fassung ist der Schutzgegenstand, das Ursprungswerk, noch erkennbar und in seinem Gesamteindruck nicht abgeändert.⁸⁷ Bestärkt wird dies durch § 23 Abs. 3 UrhG:⁸⁸ Die Segmentierung ist erforderlich, um das Korpus analysierbar zu machen, deswegen muss sie als ausschließlich technisch bedingt i. S. d. § 23 Abs. 3 UrhG eingeordnet werden.

⁸⁴ Etwas anderes gilt etwa für Bildgedichte, bei denen es auf die Position der Wörter auf der Seite ankommt. Diese sind aber nicht Gegenstand des Gutachtens.

⁸⁵ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 81.

⁸⁶ Selbst bei Texten, bei denen es auf ein Versmaß ankommt, kann das Versmaß über die späteren Annotationen beibehalten werden, da es für die Analyse möglicherweise auch relevant werden kann.

⁸⁷ Eine ähnliche Argumentation verfolgt Spindler, GRUR 2016, S. 1112, 1114.

⁸⁸ Zu § 23 Abs. 3 UrhG bereits o. S. 18 f.

Weil die Klartexte zu Zwecken der Segmentierung dupliziert werden, bleibt es insofern bei einer urheberrechtlichen Vervielfältigung, § 16 UrhG. Diese Vervielfältigung ist in Fortführung der bisherigen Argumentation sowohl durch Art. 3 Abs. 1 der DSM-RL als auch durch § 60d UrhG gerechtfertigt.

4.6.2 Änderungen, § 62 UrhG

Daneben kann das Änderungsverbot betroffen sein, denn dieses erteilt dem Urheber das Recht, zu entscheiden, wie sein Werk in Erscheinung tritt.⁸⁹ Das kann nur bei einer Sichtbarkeit für die Öffentlichkeit geschehen,⁹⁰ was in diesem Stadium, wie bereits dargelegt, noch nicht der Fall ist. Weiter gilt auch hier, dass das Änderungsverbot aufgrund der Herstellungsfreiheit von Bearbeitungen verdrängt wird.⁹¹

Kurz zusammengefasst...

Auch Segmentierungen sind urheberrechtliche Vervielfältigungen, § 16 Abs. 1 UrhG, und von § 60d UrhG erlaubt. Sofern im Einzelfall eine Änderung, § 62 UrhG, vorliegt, ist diese jedenfalls nicht dazu geeignet, die Interessen der Rechteinhaber zu beeinträchtigen, weil es nicht für die Öffentlichkeit sichtbar ist, daher ist eine Änderung auch in diesem Fall zulässig.

4.7 Bereinigung der Texte

Bei manchen Text-Mining-Projekten werden in einem weiteren Schritt sog. Stopp-Wörter aus den Texten herausgelöscht. Das sind Wörter, denen wenig oder keine semantische Bedeutung beigemessen wird, die das Analyseergebnis also verfälschen würden. In den diesem Gutachten zugrundeliegenden Anwendungsfällen wurden längere Zitate herausgelöscht. Im Entfernen fremder Inhalte ist dann keine urheberrechtlich relevante Handlung zu erkennen, wenn diese Zitate nur als fremde Schöpfung eingebracht wurden. Etwas anderes gilt aber, wenn fremde Zitate in die eigene Argumentationsstruktur aufgenommen werden oder längere Textpassagen des Werks gelöscht werden. Dann gelten die bisherigen Argumentationen, d. h. es muss entschieden werden, ob es eine wesentliche Veränderung und damit eine Bearbeitung darstellt, in Bezug auf

⁸⁹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 62 Rn. 5.

⁹⁰ Vgl. Fn. 89.

⁹¹ Vgl. dazu bereits auf S. 13 f.

1	Wir	_	PPER PPER	_	case=nom number=pl gender=* person=1	2	SUBJ	_	_
2	fragen	_	VVFIN VVFIN	_	number=pl person=1 tense=past mood=ind	0	S	_	_
3	die	_	ART ART	_	case=acc number=pl gender=* 4	DET	_	_	_
4	Arbeitsleute	_	NN NN	_	case=acc number=pl gender=* 2	OBJA	_	_	_
5	,	_	,\$ \$,	_	0	ROOT	_	_	_
6	ob	_	KOUS KOUS	_	15	KONJ	_	_	_
7	sie	_	PPER PPER	_	case=nom number=pl gender=* person=3	15	SUBJ	_	_
8	uns	_	PPER PPER	_	case=dat number=pl gender=* person=1	14	OBJD	_	_
9	nicht	_	PTKNEG PTKNEG	_	15	ADV	_	_	_
10	die	_	ART ART	_	case=acc number=sg gender=fem	11	DET	_	_
11	Spur	_	NN NN	_	case=acc number=sg gender=fem	14	OBJA	_	_
12	eines	_	ART ART	_	case=gen number=sg gender=masc	13	DET	_	_
13	Tigers	_	NN NN	_	case=gen number=sg gender=masc	11	GMOD	_	_
14	anzugeben	_	VVIZU VVIZU	_	15	OBJI	_	_	_
15	wüßten	_	VVFIN VVFIN	_	number=pl person=3 tense=past mood=subj	2	NEB	_	_
16	.	_	,\$ \$,	_	0	ROOT	_	_	_

Abbildung 2: Annotationen. Quelle: Darstellung von Melanie Andresen; Beispielsatz aus „Eine Frauenfahrt um die Welt“ von Ida Pfeiffer (1850), CC BY 4.0 International.

Änderungen gilt ebenso, dass § 62 UrhG aufgrund der Herstellungsfreiheit von Bearbeitungen verdrängt wird.⁹²

4.8 Annotationen

Im nächsten Schritt werden die Tokens oder Tokensequenzen annotiert (Abbildung 2). Die Art und Inhalte der Annotationen können von linguistischen Kategorien wie der Wortart des Tokens bis hin zu individuellen und interpretatorischen Annotationen reichen.⁹³ Generell kommt es in den digitalen Geisteswissenschaften häufig vor, dass die Annotationen händisch erfolgen. Auch in Bezug auf die Annotationen müssen Bearbeitungen, § 23 Abs. 1 UrhG, und Änderungen, § 62 UrhG, erwogen werden. Dabei ist zu beachten, dass die Annotation eines einzelnen bzw. aufgrund seiner Kürze urheberrechtlich nicht geschützten Textsegmentes nicht urheberrechtlich relevant ist, die Annotationen können jedoch die schutzfähigen Elemente des gesamten Textes betreffen bzw. dessen Gesamteindruck verändern.

4.8.1 Bearbeitungen, § 23 Abs. 1 UrhG

Gegen eine Einordnung von Annotationen als Bearbeitungen, § 23 Abs. 1 UrhG, wird argumentiert, dass die Metainformationen das Werk nur anreichern bzw. systematisieren, aber nicht den Schutzgegenstand berühren.⁹⁴ Im Falle von grammatischen Annotationen, die etwa die Wortart kategorisieren, mag dies

⁹² Ebd.

⁹³ Vgl. dazu die Beschreibung der Anwendungsfälle von XSample, S. 5 ff. sowie Andresen et al., ZfdG im Erscheinen 2022.

⁹⁴ Spindler, GRUR 2016, S. 1112, 1113 f; Spindler, ZGE 2018, S. 273, 276.

zutreffen. Interpretatorische Annotationen setzen sich aber gerade mit dem Geistesgehalt auseinander und können damit den Schutzgegenstand betreffen. Das gilt insbesondere bei fiktionaler Literatur, bei der gerade die fiktionale Erzählung als solche urheberrechtlich geschützt sein kann, wenn sie ausreichend konkretisiert wurde.⁹⁵ Dennoch stehen die Annotationen neben dem eigentlichen Werk und zeigen nur darüber hinausgehende Interpretationsansätze auf, statt das Werk selbst zu verändern. Insofern wird hier davon ausgegangen, dass Annotationen das Werk nicht bearbeiten.

4.8.2 Änderungen, § 62 UrhG

Zudem ist auch bei den Annotationen an Änderungen (§ 62 UrhG) zu denken. Diesbezüglich gilt ebenfalls das Besagte: Sie werden erst dann relevant, wenn die geänderten Werke Dritten zugänglich gemacht werden, zuvor gilt der Grundsatz der Herstellungsfreiheit von Bearbeitungen. Zudem sind die Interessen der Rechteinhaber ohne Öffentlichkeitsbezug nicht betroffen. Doch auch bei Öffentlichkeitsbezug, wie er im weiteren Forschungsprozess eintreten kann, ist zu beachten, dass nicht der Eindruck entsteht, das Werk entstamme in dieser Form vom Rechteinhaber⁹⁶ und die Interessen der Rechteinhaber nicht nennenswert betroffen sind. Zudem sind wissenschaftliche Nutzungen, insbesondere im Rahmen des TDM, gesetzlich besonders privilegiert. Bei dem TDM ist es gesetzgeberisch gerade gewollt, dass die Vorverarbeitungsschritte, auch Annotationen, erfolgen können.⁹⁷ Deswegen muss die Interessenabwägung, die im Rahmen des § 62 Abs. 1 S. 1, 2, 39 Abs. 2 UrhG erfolgt, auch dann zugunsten der Schrankenbegünstigten ausfallen.

Kurz zusammengefasst...

Mit den in TDM-Projekten in den digitalen Geisteswissenschaften üblicherweise vorgenommenen Annotationen gehen Änderungen i. S. d. § 62 UrhG einher. Gem. § 62 Abs. 1 S. 1, 2, 39 Abs. 2 UrhG muss deswegen eine Interessenabwägung stattfinden. Diese hat aber zugunsten der Schrankenbegünstigten auszufallen, wenn der Schranken Zweck der wissenschaftlichen Nutzungen im Rahmen des TDM berücksichtigt wird.

⁹⁵ Raue, ZUM 2019, S. 684, 686; Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 86 ff.

⁹⁶ Vgl. dazu bereits S. 19 ff.

⁹⁷ Die Werke werden „kategorisiert“, BT-Drs. 18/12329, S. 40. Daran sollte die neue Normierung nichts ändern BT-Drs. 19/27426, S. 87 f.

4.9 Folgeprobleme bei der Weiterverwendung veränderter oder bearbeiteter Kopien

Wenn die Abwandlungen in den Vorverarbeitungsschritten so umfangreich erfolgten, dass eine Bearbeitung oder relevante Änderung anzunehmen ist, ergeben sich Folgefragen, denn urheberrechtliche Bearbeitungen (§ 23 Abs. 1 UrhG) und deswegen auch geänderte Werkexemplare (§ 62 UrhG) dürfen nur hergestellt, nicht aber weiterverwertet werden.

4.9.1 Umgang mit dem Bearbeitungsrecht, § 23 UrhG

Die deutschen sowie die europäischen Normen basieren auf der Annahme, dass umfangreiche Vorverarbeitungen erfolgen, d. h. auch davon, dass die Ursprungsmaterialien in verschiedenen Schritten vorverarbeitet werden. Aus diesem Grund müssen die §§ 44b, 60d UrhG dergestalt ausgelegt werden, dass nicht nur identische, sondern auch bearbeitete Vervielfältigungen im Zuge der weiteren Forschungsarbeiten weiter verwertet werden dürfen. Diese Auslegung ist aus unionsrechtlicher Perspektive sowohl möglich als auch erforderlich, denn das Bearbeitungsrecht, wie es im deutschen Recht geregelt ist, existiert im europäischen Kontext nicht – das Unionsrecht geht stattdessen von abgewandelten Vervielfältigungen aus.⁹⁸ Diese müssen im weiteren Fortgang auch im Rahmen der Art. 3, 4 DSM-RL weiter vervielfältigt werden dürfen. Auch die vom deutschen Gesetzgeber auf Art. 5 Abs. 3 lit. a InfoSoc-RL gestützten öffentlichen Zugänglichmachungen für gemeinsame Forschungsarbeiten oder wissenschaftliche Überprüfungen, die an späterer Stelle thematisiert werden, können insofern ermöglicht werden.

4.9.2 Umgang mit dem Änderungsverbot, § 62 UrhG

In Bezug auf das Änderungsverbot, § 62 UrhG, das auf dem nationalstaatlich geregelten Urheberpersönlichkeitsrecht basiert,⁹⁹ ist die Problemstellung

⁹⁸ Leenen in Wandtke/Bullinger, UrhG, Art. 2 InfoSoc-RL Rn. 7, Art. 1 InfoSoc-RL Rn. 2; v. Ungern-Sternberg in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 15 Rn. 32 ff; 242 ff; Grünberger, ZUM 2015, S. 273, 275; BGH, Urt. v. 16. 5. 2013 – I ZR 28/12, Beuys-Aktion = GRUR 2014, 65; BGH, Urt. v. 29. 4. 2010 – I ZR 69/08, Vorschaubilder = GRUR 2010, 628, S. 629 Rn. 17; EuGH, ECLI:EU:C:2015:27, Allposters/Pictoright = GRUR 2015, 256, S. 257 Rn. 27.

⁹⁹ Peukert in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 62 Rn. 1; Bullinger in Wandtke/Bullinger, UrhG, § 62 Rn. 2; Peukert in Loewenheim UrhR-HdB, § 15 Rn. 1; BT-Drs. IV/270, S. 76.

schwieriger zu lösen. Systematisch sieht § 62 UrhG zugunsten mancher urheberrechtlichen Schranken explizite Änderungserlaubnisse vor. Ansonsten verweist die Norm auf § 39 Abs. 2 UrhG, der solche Änderungen zulässt, die der Rechteinhaber nach Treu und Glauben nicht verwehren darf. Maßstäbe dieser Abwägung sind insbesondere der gesetzlich anerkannte Nutzungszweck der urheberrechtlichen Schranke.¹⁰⁰ Ergänzend wird teilweise auf die Kriterien des Entstellungsverbots, § 14 UrhG, zurückgegriffen.¹⁰¹ In dessen Rahmen sind das „Bestands- und Integritätsinteresse“¹⁰² des Urhebers mit den Nutzungsinteressen abzuwägen.¹⁰³ Dabei kommt es natürlich immer auf den jeweiligen Einzelfall an. Eine Gefährdung der Interessen der Rechteinhaber kommt allerdings nur in Betracht, wenn das entstellte Werk der Öffentlichkeit sichtbar gemacht wird,¹⁰⁴ denn das Änderungsverbot soll den Urheber davor schützen, dass sein Werk in der Öffentlichkeit mit Änderungen präsentiert wird, die dem Urheber zugerechnet werden.¹⁰⁵ In dem Stadium des Forschungsprozesses, in dem (fehlerhafte) Digitalisierungen erstellt werden, verbleibt der Scan innerhalb der Forschungsgruppe. Dann liegt nur in den seltensten Fällen eine Öffentlichkeit vor.¹⁰⁶ Falls dem doch so sein sollte, besteht aufgrund des Charakters der TDM-Analyse bzw. der Kenntnis um den umfangreichen Vorverarbeitungsprozess jedenfalls die volle Kenntnis darüber, dass es sich bei den jeweiligen Vervielfältigungsstücken nicht um das Werk in seiner vom Urheber bezweckten Gestalt handelt, sondern um eine zu Forschungsdaten transformierte Fassung. Die Vorverarbeitung ist also als Ergebnis wissenschaftlicher Forschung erkennbar. Deswegen sind die Vorverarbeitungsschritte schon nicht dazu geeignet, die Interessen der Rechteinhaber zu gefährden. Zudem ist wertungsmäßig vorzubringen, dass die Nutzung gerade im Rahmen einer unionsrechtlich verpflichtenden urheberrechtlichen Schranke erfolgt. Insofern liegt in der Weiterverarbeitung von geänderten Werken kein Verstoß gegen § 62 UrhG. Ein ähnliches Ergebnis erlangt man mittels eines Erst-Recht-Schlusses: Wenn schon Bearbeitungen ausgenommen sind, müssen es erst recht Änderungen sein.¹⁰⁷

¹⁰⁰ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 62 Rn. 12.

¹⁰¹ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 39 Rn. 16.

¹⁰² Peukert in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 14 Rn. 26.

¹⁰³ Ebd.; BGH, Urt. v. 29.04.1970 – I ZR 30/69, Maske in Blau = GRUR 1971, 35, S. 37.

¹⁰⁴ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 14 Rn. 8; Dustmann in Fromm/Nordemann, UrhG, § 14 Rn. 16; so auch schon das RG, Urt. v. 8.6.1912 – I 382/11, Felseneiland mit Sirenen = RGZ 79, 397, S. 402.

¹⁰⁵ Peukert in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 14 Rn. 1.

¹⁰⁶ Ebenso Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 12.

¹⁰⁷ So Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 23 Rn. 85.

Kurz zusammengefasst...

Die rechtlichen Probleme, die daraus ergeben, dass bearbeitete oder geänderte Fassungen des Werks weitergenutzt werden, lassen sich mittels einer erweiterten Auslegung der §§ 44b, 60d UrhG sowie der Interessenabwägung des § 62 UrhG lösen. Insofern sind in dem thematisierten Umfang im geltenden Recht auch die Veränderung des äußeren Erscheinungsbildes und kleinere Veränderungen des Werks möglich.

4.10 Pflicht zur Quellenangabe, § 63 UrhG

Urheber haben gem. §§ 13, 63 UrhG das Recht, als solche anerkannt zu werden. Das bedeutet konkret, dass der jeweilige Urheber (Vor- und Zuname), der vollständige Werktitel, Auflage und Erscheinungsjahr sowie – wenn ganze Werke enthalten sind – der Name des Verlags (§ 63 Abs. 1 S. 2 UrhG) deutlich gemacht werden müssen.¹⁰⁸ Insgesamt muss die Fundstelle erkennbar sein, d. h. bei Publikumszeitschriften und Fachzeitschriften muss auch deren Name genannt werden, § 63 Abs. 3 UrhG. Bei Webseiten ist das die URL.¹⁰⁹ Darüber hinaus sieht § 63 Abs. 1 S. 2 UrhG vor, dass deutlich gemacht werden muss, ob Kürzungen oder Änderungen vorgenommen wurden. Gerade in Bezug auf die vollzogenen Vorverarbeitungsschritte bietet es sich allerdings an, diese in gebotener Kürze separat von den analysierten Dateien selbst zu dokumentieren. Eine an das Korpus angefügte Datei, die die Schritte enthält, ist insofern zu bevorzugen.

Seit der Umsetzung der DSM-RL gilt § 63 Abs. 1 UrhG nicht mehr bei Vervielfältigungen im Rahmen des § 60d UrhG.¹¹⁰ Insofern sind jedenfalls die Vervielfältigungen möglich, ohne dabei die Quellenangabe zu beachten. Etwas anderes gilt allerdings für die öffentliche Zugänglichmachung, z. B. bei gemeinsamen Forschungsarbeiten,¹¹¹ bei denen gem. § 63 Abs. 2 UrhG auch bei Nutzungen nach § 60d UrhG die Quelle anzugeben ist. Diesbezüglich ist eine Erleichterung der Quellenangabe gem. § 63 Abs. 2 S. 2 UrhG in Betracht

¹⁰⁸ Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 63 Rn. 11 ff; Spindler in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 63 Rn. 13 f.

¹⁰⁹ Spindler in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 63 Rn. 14.

¹¹⁰ Die Anforderungen des § 63 UrhG wurde schon in Bezug auf das UrhWissG kritisiert, vgl. Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, § 60d Rn. 9; Hegemann/Naldony in Hoeren/Sieber/Holznapel, Multimediarecht Hdb, § 60d Rn. 139; Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 19; Spindler, ZGE 2018, S. 273, 283.

¹¹¹ Vgl. dazu S. 29.

zu ziehen, nach dem auf die Angabe der Quelle verzichtet werden darf, wenn sie nicht möglich ist. Das diese Erleichterung anwendbar ist, wird allerdings als unwahrscheinlich erachtet, gerade weil die Nennung der erforderlichen Angaben rein technisch möglich ist.¹¹² Jedenfalls erleichtert es die Forschungsarbeit erheblich, wenn bereits zu Beginn die Quellen angegeben werden, um spätere zusätzliche Arbeiten zu vermeiden. Das ist zumeist ohnehin praktiziert.

4.11 Das Korpus

Aus den erläuterten Schritten geht nun das sog. Korpus hervor. Es enthält die segmentierte und annotierte Fassung, also den eigentlichen Analysegegenstand, zumeist werden aber jedenfalls auch die OCR-Digitalisate aufbewahrt, die noch zur Nachprüfung der Abläufe relevant werden können.

4.11.1 Schutz der Korpora

Es liegt nahe, dass derartige Korpora als Datenbanken urheberrechtlich geschützt sind, etwa als Datenbankwerke, § 4 Abs. 2, 2 Abs. 2 UrhG, oder Datenbanken, §§ 87a ff. UrhG. Grundsätzlich können Datenbanken auch die Voraussetzungen beider Schutzrechte erfüllen.¹¹³

4.11.1.1 Sammel- und Datenbankwerke, §§ 4 Abs. 1, 2, 2 Abs. 2 UrhG

Das UrhG schützt Sammlungen u. a. von Werken, die persönliche bzw. eigene geistige Schöpfungen sind, als Sammelwerke, § 4 Abs. 1 UrhG, und als Datenbankwerke, § 4 Abs. 2 UrhG. Sammelwerke sind dabei „*Sammlungen von Werken, Daten oder anderen unabhängigen Elementen, die aufgrund der Auswahl oder Anordnung der Elemente eine persönliche geistige Schöpfung sind*“. Unter Datenbankwerken hingegen versteht man Sammelwerke, deren Elemente „*systematisch oder methodisch angeordnet und einzeln mit Hilfe elektronischer Mittel oder auf andere Weise zugänglich sind*“. Kennzeichnend ist jedoch, dass das Kriterium der Unabhängigkeit der Elemente universell gilt, d. h. auch die enthaltenen Werke oder Daten müssen voneinander unabhängig sein,¹¹⁴ das bedeutet, dass bzgl. der einzelnen Elemente ein „*hinreichender*

¹¹² Schulze in Dreier/Schulze, UrhG, 7. Auflage 2022, § 63 Rn. 25; Spindler in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 63 Rn. 18a. Anderer Auffassung ist Raue, ZUM 2021, S. 793, 799 der davon ausgeht, dass die Quellenangabe technisch unmöglich ist.

¹¹³ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 4 Rn. 3; BGH, Urt. v. 24. 5. 2007 – I ZR 130/04, Gedichttitelliste I = GRUR 2007, 685.

¹¹⁴ Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 4 Rn. 18.

Informationswert“ besteht.¹¹⁵ Die vom Datenbankrecht unter Schutz gestellte Leistung liegt in der Herstellung eigener Verbindungen zwischen den Elementen. Die Schöpfung liegt dabei genauer in der Originalität bzw. der Individualität der Auswahl oder Anordnung des Inhalts.¹¹⁶ Der Schutz von Korpus und enthaltenen Elementen besteht unabhängig voneinander.¹¹⁷ Zum Beispiel wurde von der Rechtsprechung bei einer Gedichttitelliste der 1100 wichtigsten Gedichten aus der Zeit zwischen 1730 und 1900 samt Metadaten angenommen, dass es sich um ein Datenbankwerk handelt.¹¹⁸

Grundsätzlich kommen bei Korpora aus den digitalen Geisteswissenschaften beide Schutzrechte in Betracht, wobei der Schutz von Datenbankwerken ein geringeres Maß an Individualität der Sammlung erfordert, es genügt, dass ein bestimmtes Schema und eine bestimmte Herangehensweise verfolgt werden und das System maschinell abrufbar ist.¹¹⁹ Aufgrund der Anwendbarkeit des unionsrechtlichen Werkbegriffs im Bereich des harmonisierten Datenbankwerkschutzes, Art. 3 Abs. 1 Datenbank-RL, ist die dabei notwendige Schöpfungshöhe im Bereich der Datenbankwerke zusätzlich niedriger angesiedelt.¹²⁰ In Zweifel zu ziehen ist die Unabhängigkeit der Elemente bei Korpora aus den digitalen Geisteswissenschaften zwar deswegen, weil das Korpus als Ganzes analysiert werden soll und die Annotationen in ihrer Gesamtheit Muster oder Strukturen hervorbringen soll. Allerdings wird gerade in den digitalen Literaturwissenschaften umfangreich annotiert (Abbildung 2), weswegen eine Unabhängigkeit anzunehmen ist. Technisch ist es auch möglich, auf die Korpussegmente einzeln zuzugreifen, wie es der Datenbankwerkschutz erfordert. Gegen eine Schutzfähigkeit kann wiederum sprechen, dass die Art und Weise der Zusammenstellung der Korpora technisch bedingt sein kann, was eine Individualität ausschließt. Allerdings fließt gerade im geisteswissenschaftlichen Bereich ein großer gedanklicher Aufwand in die Gestaltung des Forschungskonzeptes und die Auswahl der analysierten Literatur, was eine hinreichende Individualität der Konzeption begründen kann. Auch die hinzugefügten

¹¹⁵ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 4 Rn. 10; zu Datenbanken nach der Datenbank-RL der EuGH, ECLI:EU:C:2015:735, Esterbauer = GRUR 2015, 1187, S. 1188 Rn. 17; EuGH, ECLI:EU:C:2004:696, Fixtures-Fußballspielpläne II = GRUR 2005, 254, S. 255 Rn. 29; Hacker, ZGE 2020, S. 239, 1029.

¹¹⁶ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 4 Rn. 11, 19; Leistner in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 4 Rn. 19.

¹¹⁷ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 4 Rn. 4.

¹¹⁸ BGH, Urt. v. 24. 5. 2007 – I ZR 130/04, Gedichttitelliste I = GRUR 2007, 685, S. 687 Rn. 19 ff.

¹¹⁹ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 2 Rn. 17.

¹²⁰ Vgl. dazu bereits o. S. 9 ff.

Metadaten und Annotationen stützen die gedankliche Konzeption, die ein eigenes System zur Erforschung eines Phänomens darstellt.

Insofern können Korpora der digitalen Geisteswissenschaften sowohl als Sammelwerke, § 4 Abs. 1, § 2 Abs. 2 UrhG, als auch als Datenbankwerke, §§ 4 Abs. 2, § 2 Abs. 2 UrhG, geschützt sein.¹²¹

4.11.1.2 Datenbanken, §§ 87a ff. UrhG

Das Datenbankherstellerrecht, §§ 87a ff. UrhG, setzt nicht voraus, dass eine persönliche geistige Schöpfung vorliegt. Das Datenbankherstellerrecht erfasst Sammlungen unabhängiger Elemente, deren Beschaffung, Überprüfung oder Darstellung eine nach Art oder Umgang wesentliche Investition erfordert. Geschützt wird unter diesen Voraussetzungen das bearbeitete und durchsuchbare Datenmaterial, nicht aber Rohdaten selbst.¹²²

Wie bereits thematisiert wurde, können die Korpuselemente voneinander unabhängig sein. Die Sammlung des Datenbankinhalts erfordert in diesem Sinne eine systematische oder methodische Anordnung, es muss also ein System zugrunde liegen.¹²³ Das ein TDM-Korpus in den digitalen Geisteswissenschaften systematisch und methodisch angeordnet ist, wird im in Abbildung 2 dargestellten Korpusauszug deutlich, gerade die Annotationen erfolgen nach einem festgelegten Konzept, das für die maschinelle Analysierbarkeit bzw. die Aussagekraft der Analyseergebnisse zwingend notwendig ist. Unter der Beschaffung ist die „*sichtende, beobachtende und auswertende*“ Tätigkeit des Datenbankherstellers zu verstehen.¹²⁴ Es sind aber auch Aufwendungen für die Darstellung des Datenbankinhalts im Rahmen des § 87a UrhG zu berücksichtigen, d. h. solche Aufwendungen, die der Anordnung, Organisation, Zugänglichkeit und Aktualität dienen.¹²⁵ Dazu zählen auch explizit Kosten der

¹²¹ So bereits in Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 197.

¹²² Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 87a Rn. 7.

¹²³ Ebd.

¹²⁴ Vogel in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 87a Rn. 44; u. a. EuGH, ECLI:EU:C:2004:695, British Horseracing Board = GRUR 2005, 244, S. 247 Rn. 31 f; EuGH, ECLI:EU:C:2004:696, Fixtures-Fußballspielpläne I = GRUR 2005, 252; BGH, Teilurt. v. 30. 4. 2009 – I ZR 191/05, Elektronischer Zolltarif = GRUR 2009, 852, S. 853 Rn. 23 Dem liegt die Erwägung zugrunde, dass gerade die Datenbankherstellung, nicht dagegen die Datenerzeugung unterstützt werden sollte, EuGH, ECLI:EU:C:2004:696, Fixtures-Fußballspielpläne I = GRUR 2005, 252, S. 253 Rn. 24.

¹²⁵ Vogel in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 87a Rn. 45; EuGH, ECLI:EU:C:2004:696, Fixtures-Fußballspielpläne I = GRUR 2005,

Datenaufbereitung.¹²⁶ In den Anwendungsfällen fließt der größte Aufwand in die Aufbereitung der Forschungsdaten. Insofern können die Korpora auch als Datenbanken gem. § 87a ff. UrhG eingeordnet werden.¹²⁷

4.11.2 Rechteinhaberschaft

Die Nutzungsrechte an den Korpora stehen in den meisten Fällen auf Rechtsfolgenseite den Arbeitgebern zu: Urheber ist bzw. sind zwar stets die Person oder Personen (Miturheber, § 8 UrhG), die die Auswahl getroffen hat oder haben.¹²⁸ Sofern das Korpus als Datenbankwerk, §§ 4 Abs. 2, 2 Abs. 2 UrhG, im Rahmen eines abhängigen Beschäftigungsverhältnisses geschaffen wurde, gilt § 43 UrhG, der besagt, dass die Vorschriften zum Urhebervertragsrecht auch in Arbeits- oder Dienstverhältnissen gelten, soweit sich aus dem jeweiligen Vertragsverhältnis nichts anderes ergibt. Demzufolge muss bei Werken, sofern sie in Erfüllung der arbeits- oder dienstvertraglichen Verpflichtungen entstanden sind (d. h. bei sog. Pflichtwerken), geprüft werden, ob das Vertragsverhältnis ergibt, dass die Nutzungsrechte dem Arbeit- oder Dienstgeber zustehen.¹²⁹ In dem Fall besteht u. U. eine Verpflichtung zur Einräumung von Nutzungsrechten, die in der Regel spätestens mit der Übergabe des Werkstücks erfolgt.¹³⁰ Zusätzlich sind die Besonderheiten der Wissenschaftsfreiheit, Art. 5 Abs. 3 GG, zu beachten. Außerhalb der sog. Pflichtforschung ist die Rechtslage anders zu beurteilen.¹³¹

Das Datenbankherstellerrecht steht im Gegensatz zum Datenbankurheberrecht originär demjenigen zu, der die Investition getätigt hat, § 87a Abs. 2 UrhG. Das ist derjenige, der „die Initiative ergreift und das Investitionsrisiko trägt“, auch

252, S. 253 Rn. 27; EuGH, ECLI:EU:C:2004:696, Fixtures-Fußballspielpläne II = GRUR 2005, 254, S. 256 Rn. 43.

¹²⁶ Vogel in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 87a Rn. 45; tendenziell auch Gaster in Hoeren/Sieber/Holzngel, Multimediarecht Hdb, Teil 7.6 Rn. 86 ff.

¹²⁷ So ebenso bereits in Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 197.

¹²⁸ BGH, Urt. v. 24. 5. 2007 – I ZR 130/04, Gedichttitelliste I = GRUR 2007, 685, S. 687 Rn. 23.

¹²⁹ Ausgenommen sind aber zeitlich vor- oder nachgelagert entstandene Werke sowie außervertragliche Werke, Rojahn/Frank in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 43 Rn. 21 ff.; Rieger, Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 124 f.

¹³⁰ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 43 Rn. 18 ff; Rojahn/Frank in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 43 Rn. 37 ff; Rieger, Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 124 f; BGH, Urt. v. 22.02.1974 – I ZR 128/72, Hummelrechte = GRUR 1974, 480, S. 483.

¹³¹ Rojahn/Frank in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 43 Rn. 131 ff; Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, § 43 Rn. 12.

wenn er diese die Datenbank nicht konzipiert hat.¹³² Nicht als Investitionen betrachtet werden in diesem Kontext die getätigten Arbeitsleistungen.¹³³ Rechteinhaber sind demnach nicht notwendigerweise die Arbeitnehmer, sondern in aller Regel die hinter den Forschenden stehende natürliche oder juristische Person,¹³⁴ z. B. der Projektträger oder Arbeitgeber.

4.11.3 Rechtsverletzungen durch Integrierung der aufbereiteten Werke inein Forschungskorpus

In beiden Fällen, ob das Korpus als Datenbankwerk oder als Datenbank geschützt ist, bleiben die Rechte an den in den Datenbanken enthaltenen Schutzgegenständen erhalten, solange die Werke noch wiedererkennbar sind.¹³⁵ Offen ist deswegen, ob in der Eingliederung einzelner annotierter Werke gemeinsam mit anderen annotierten Werken in ein Gesamt-Korpus eine Rechtsverletzung liegt.

Wenn die annotierten Fassungen kopiert werden, um sie zu einem Korpus zusammenzufügen, werden sie jedenfalls gem. § 16 Abs. 1 UrhG vervielfältigt. Das ist gem. §§ 60d Abs. 1, 2, 3, 44b Abs. 1 UrhG rechtmäßig. Weil die Werke aber auch u. U. in einen anderen Sachzusammenhang gestellt werden, muss auch geprüft werden, ob sie damit i. S. d. § 23 UrhG bearbeitet werden, denn dadurch kann auch eine Bearbeitung begründet werden.¹³⁶ Die Erstellung eines Korpus ist urheberrechtlich freigestellt, wie bereits dargelegt, sind aber weitere Nutzungen ins Auge zu fassen. Mit der gesetzgeberisch bezweckten Korpuserstellung¹³⁷ geht gerade einher, dass viele Werke oder Schutzgegenstände aneinandergesetzt werden. Insofern gilt auch hier, dass die urheberrechtliche Schranke des § 60d UrhG entgegen des Wortlauts, der nur Vervielfältigungen erlaubt, auch für Bearbeitungen bzw. deren Weiterverwendung gilt.

Weiter kann die Korpuseingliederung auch als Änderung gem. § 62 Abs. 1 UrhG zu bewerten sein. Doch auch in diesen Fällen muss die Interessenabwägung nach §§ 62 Abs. 1 S. 2, 39 Abs. 2 UrhG zugunsten des Schrankenbegünstigten ausfallen: Das Korpus wird in diesem Schritt, wie in den

¹³² Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 87a Rn. 19; Richtlinie (EU) 96/6/EG, Datenbank-RL Erw. 41.

¹³³ Rieger, Der rechtliche Schutz wissenschaftlicher Datenbanken, S. 206; Richtlinie (EU) 96/6/EG, Datenbank-RL Erw. 41.

¹³⁴ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 43 Rn. 19 f.

¹³⁵ Das Kriterium der Wiedererkennbarkeit wurde vom EuGH geprägt, s. EuGH, ECLI:EU:C:2019:624, Pelham/Hütter [Metall auf Metall III] = GRUR 2019, 929, S. 931 Rn. 31, 36 f. und vom deutschen Gesetzgeber adaptiert, BT-Drs. 19/27426, S. 78; Schack, GRUR 2021, S. 904, 906; ausführlich Grisse/Kaiser, ZUM 2021, S. 401.

¹³⁶ Loewenheim in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 23 Rn. 6.

¹³⁷ BT-Drs. 18/12329, S. 40 f; BT-Drs. 19/27426, S. 87 f.

bisherigen Schritten auch, nur in seltenen Fällen einer Öffentlichkeit im Sinne der Allgemeinheit, zugänglich gemacht. Überdies entstünde, selbst wenn im Einzelfall eine Öffentlichkeit anzunehmen wäre, nicht der Eindruck, das abgewandelte Werk stamme vom Autor. Es ist vielmehr intendiert, dass die Vorverarbeitung als Ergebnis wissenschaftlicher Forschung erkennbar ist. Zudem ist wertungsmäßig vorzubringen, dass die Integrierung in das Korpus rechtmäßig im Rahmen einer urheberrechtlichen Schranke erfolgt. Deswegen fiele auch eine Interessenabwägung, wie sie im Rahmen der §§ 62 Abs. 1 S. 2, 39 Abs. 2 UrhG sowie im Rahmen des § 14 UrhG erfolgt, zugunsten der Wissenschaft aus.

Kurz zusammengefasst...

- TDM-Korpora können als Datenbanken gem. § 87a ff. UrhG sowie als Datenbankwerke gem. §§ 4 Abs. 2, 2 Abs. 2 UrhG geschützt sein. Der Schutz der in den Korpora enthaltenen Werke bleibt neben dem Schutz der Korpora selbst bestehen.
- Beim Zusammenstellen der Korpora können dadurch Bearbeitungen (§ 23 UrhG) erfolgen, dass die Werke in einen anderen Sachzusammenhang gestellt werden. Die Herstellung dieser Bearbeitung ist frei und auch die weitere Nutzung ist wegen einer erweiterten Auslegung der §§ 60d, 44b UrhG rechtmäßig. Zudem kann eine urheberrechtliche Änderung, § 62 UrhG, vorliegen, die aufgrund der gesetzgeberischen Zwecksetzung und überwiegender wissenschaftlicher Interessen als zulässig einzuordnen ist.

4.12 Gemeinsame Forschungsarbeiten

§ 60d Abs. 4 Nr. 1 UrhG erfasst die Möglichkeit, die Vervielfältigungen, d. h. die Korpora, nicht allerdings die unveränderten Ursprungsdaten,¹³⁸ einem bestimmt abgegrenzten Personenkreis für die gemeinsame wissenschaftliche Forschung öffentlich zugänglich zu machen. Damit bezieht der Absatz sich auf das Recht der öffentlichen Zugänglichmachung gem. § 19a UrhG, schränkt die Erlaubnis aber etwas weiter ein, indem sie keine Zugänglichmachung an die Gesamt-Öffentlichkeit im Sinne einer Allgemeinheit erlaubt, sondern nur an bestimmt abgegrenzte Teile der Öffentlichkeit. Eine Öffentlichkeit ist jedoch nur dann gegeben, wenn keine persönliche Beziehung zu den Adressaten

¹³⁸ Anton in Spindler/Schuster, Recht der elektronischen Medien, § 60d Rn. 13. Das gilt nach dem Wortlaut des § 60d UrhG n. F. weiterhin.

besteht und es sich um eine unbestimmte Anzahl sowie eine „*ziemlich große Zahl von Personen*“ handelt.¹³⁹ Jedenfalls bei kleinen Forschungsgruppen liegt zumeist keine Öffentlichkeit vor, weswegen die Erlaubnis häufig nicht erforderlich ist.¹⁴⁰ Die öffentliche Zugänglichmachung ist nach Abschluss der Forschungsarbeiten zu beenden, § 60d Abs. 4 S. 2 UrhG. Das bedeutet, dass das Korpus, sobald die Forschungsarbeiten beendet sind, nicht mehr so zur Verfügung stehen dürfen, dass Mitglieder der Öffentlichkeit, wie sie bei gemeinsamen Forschungsarbeiten in der Regel nicht betroffen ist, sie an Orten und zu Zeiten ihrer Wahl abrufen können. Sie müssen also z. B. von gemeinsamen Servern oder Clouds offline genommen werden. Wann das der Fall ist, hängt von der Themenauswahl ab, denn diese definiert gerade das Forschungsziel, unterliegt aber der Wissenschaftsfreiheit und damit innerhalb gewisser Grenzen der Einschätzungsprärogative der Forschenden.¹⁴¹

4.13 Rechtliche Bewertung der TDM-Analyse im engeren Sinne

Die TDM-Analyse im engeren Sinne ist urheberrechtlich weitaus weniger problematisch als die thematisierte Korpuserstellung. Zur eigentlichen Analyse werden in der aktuellen Forschungspraxis verschiedene Methoden verwendet. Die Analysemethoden sind wenig bis gar nicht standardisiert und finden häufig versuchsweise Anwendung. Es finden teilweise bereits vorhandene bzw. verfügbare Algorithmen Anwendung. Bei der Analyse des Korpus kann es zu temporären Vervielfältigungen im Arbeitsspeicher kommen, die von § 44a UrhG gedeckt und daher rechtmäßig sind.¹⁴² Davon abgesehen ist die Datenanalyse selbst nach allgemeiner Auffassung urheberrechtlich nicht relevant.¹⁴³ Sie muss allerdings Ziel des Gesamtprozesses sein und entsprechend Art. 2 Nr. 2 DSM-RL sowie den §§ 44b, 60d UrhG automatisiert erfolgen. Eine rein manuelle Analyse ist hingegen nicht erfasst: Die Gesetzesbegründung zu

¹³⁹ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 15 Rn. 38. v. Ungern-Sternberg in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 16 Rn. 72 ff. m. w. N; EuGH, ECLI:EU:C:2014:76, Nils Svensson ua/Retriever Sverige = GRUR 2014, 360, S. 361 Rn. 21.

¹⁴⁰ Vgl. o. Fn. 107.

¹⁴¹ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60d Rn. 12; Stieper in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 18; Raue, CR 2017, S. 656, 658 f; Raue, ZUM 2021, S. 793, 799; BVerfG, Urt. v. 29.5.1973 – 1 BvR 424/71 u. 325/72 = NJW 1973, 1176.

¹⁴² Raue, ZUM 2021, S. 793, 795.

¹⁴³ Stieper in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 11; Bullinger in Wandtke/Bullinger, UrhG, § 60d Rn. 15; BT-Drs. 18/12329, S. 40; DSM-RL Erw. 9.

§ 60d UrhG a. F. stellt diesbezüglich klar, dass „*die Inhalte nicht nur rein manuell verarbeitet werden dürfen*“.¹⁴⁴ Es ist nicht anzunehmen, dass die Gesetzesreform daran etwas ändern sollte.

Kurz zusammengefasst...

Die eigentliche Datenanalyse ist im Rahmen des TDM urheberrechtlich frei.

4.14 Vergütung

Streuend der Vorgabe der DSM-RL¹⁴⁵ sind Vervielfältigungen, die im Rahmen des TDM erfolgen, nicht vergütungspflichtig.

4.15 Publikation der Forschungsergebnisse

Die Publikationsmöglichkeiten sehen sich in den digitalen Geisteswissenschaften erheblichen Einschränkungen ausgesetzt, wenn mit TDM gearbeitet wird, denn die Überzeugungskraft wissenschaftlicher Publikationen wird dadurch gestärkt, dass sie Auszüge der Forschungsdaten enthalten. Die Nutzbarkeit von Korpusauszügen in wissenschaftlichen Publikationen begegnet aber in der Praxis erheblichen Rechtsunsicherheiten. Im Folgenden wird geprüft, ob jedenfalls Teilen der Korpora urheberrechtlich in die Publikationen der Forschungsergebnisse integriert werden können.

Klassischerweise kommt die **Zitatschranke** zum Tragen, § 51 UrhG, wenn im Rahmen der wissenschaftlichen Argumentation fremde Werke oder Anteile übernommen werden sollen. Diese Zitatschranke soll es ermöglichen, zur Erläuterung des Inhalts oder auch zu Zwecken der Kritik oder Rezension Textpassagen in neue Werke aufzunehmen. Die Aufnahme in ein neues Werk wäre bei wissenschaftlichen Publikationen zumeist gegeben. Problematisch ist allerdings, dass nicht ein Ausschnitt des Originalwerks, sondern in Korpusauszug genutzt wird, der Auszüge des Werks in einer stark abgewandelten Form enthält.

Die DSM-RL verweist in Bezug auf Nutzungen außerhalb des TDM ausdrücklich darauf, dass die Begutachtung im wissenschaftlichen Diskurs weiterhin unter Art. 5 Abs. 3 lit. a InfoSoc-RL fallen solle.¹⁴⁶ Insofern geht das EU-Recht jedenfalls davon aus, dass die Korpora weiterverwendet werden können, obwohl sie urheberrechtlich geschützte und vielfach abgewandelte

¹⁴⁴ BT-Drs. 18/12329, S. 41.

¹⁴⁵ DSM-RL Erw. 17.

¹⁴⁶ DSM-RL Erw. 15.

Daten enthalten. Infolge ist es auch grundsätzlich denkbar, dass auch Zitate aus den Korpora entnommen werden dürfen. Dazu ist die Zitatschranke in ihrem Wortlaut, der auf ein veröffentlichtes Werk abstellt, auf Bearbeitungen oder Vervielfältigungen zu erweitern. Das wird auch dadurch gestützt, dass die Zitatschranke auf Art. 5 Abs. 3 lit. d InfoSoc-RL beruht, also die Umsetzung einer Unionsbestimmung ist. Das Unionsrecht nimmt, wie bereits dargestellt,¹⁴⁷ keine Unterscheidung zwischen Vervielfältigung und Bearbeitung vor. Deswegen ist Art. 5 Abs. 3 lit. d InfoSoc-RL so auszulegen, dass sie auch die Übernahme von abgewandelten Vervielfältigungen im unionsrechtlichen Sinne erfasst. Das muss auch für die nationale Umsetzung gelten.

Die Übernahme ließe sich auf § 51 Nr. 2 UrhG zu stützen, das **wissenschaftliche Kleinzitat**, § 51 Nr. 2 UrhG das es erlaubt, Stellen eines Werks nach der Veröffentlichung in ein selbstständiges Sprachwerk aufzunehmen und zu diesem Zweck u. a. zu vervielfältigen, § 16 Abs. 1 UrhG, und öffentlich zugänglich zu machen, § 19a UrhG. Der Umfang ist tendenziell restriktiv auszulegen und im Sinne von „*kleinen Ausschnitten*“ zu verstehen.¹⁴⁸ Das bemisst sich auch nach dem Umfang des Sprachwerks, in das die Ausschnitte aufgenommen werden, d. h. bei sehr langen Sprachwerken oder zur Erläuterung besonders relevanter Textpassagen sind auch längere Übernahmen möglich.¹⁴⁹ Das **wissenschaftliche Großzitat**, § 51 Nr. 1 UrhG, erfasst dagegen nur die Aufnahme einzelner Werke in wissenschaftliche Werke. Bei den gesamten TDM-Korpora, die i. d. R. aus vielen Werken bestehen, ist diese Grenze in jedem Fall überschritten. Kleinere Korpusauszüge dürfen jedoch genutzt werden. Im Rahmen der Zitatschranke ist einschränkend die **Pflicht zur Quellenangabe** zu beachten, das gilt gem. §§ 63 Abs. 1, 2 UrhG sowohl bei Vervielfältigungen aus auch bei öffentlichen Zugänglichmachungen.

Zudem kann das **Änderungsverbot** gem. § 62 UrhG in diesem Fall ein Zitat verhindern. Im forschungsimernen Bereich lässt sich stets argumentieren, dass es die Interessen der Urheber nicht beeinträchtigt, wenn ein geändertes Werk innerhalb einer Forschungsgruppe verbleibt. Bei einer Publikation soll es aber gerade, jedenfalls anteilig, der Öffentlichkeit zugänglich gemacht werden. In diesem Fall müssen die Interessen der Wissenschaft mit denjenigen der Rechteinhaber abgewogen werden. Dazu kann ebenso argumentiert werden, dass bei der Übernahme von Korpuspassagen zum einen nur kurze Ausschnitte genutzt werden dürfen, diese aus einer durch unionsrechtlich verpflichtende Normen (Art. 3 DSM-RL) erlaubten Forschung stammen und zum anderen nicht der Eindruck entsteht, das Werk entstamme in dieser Form von dem betreffenden Urheber. Insofern steht die Interessenlage der Übernahme kurzer Korpuspassagen grundsätzlich nicht entgegen.

¹⁴⁷ Vgl. S. 23.

¹⁴⁸ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 51 Rn. 14; Spindler in Schricker/Loewenheim, UrhG, 5. Auflage 2017, § 51 Rn. 80 ff.

¹⁴⁹ Ebd.

Kurz zusammengefasst...

Die Zitatschranke kann grundsätzlich auf die Aufnahme kurze Korpuspassagen in Forschungspublikationen angewandt werden. Dabei muss die Quelle genannt werden.

KAPITEL 5

Umgang mit den Korpora nach Abschluss der Forschungsarbeiten

Neben den Unsicherheiten, die die Erlaubnis des Gesamtprozesses des TDM inklusive der Vorverarbeitungsschritte betreffen, besteht eine große Rechtsunsicherheit auch bzgl. der Abläufe, die sich an den Abschluss der eigentlichen Forschungsarbeiten anschließen, wie die Lagerung der Forschungsdaten, die wissenschaftliche Überprüfbarkeit und die Frage, ob auf die Korpora zu Zwecken von Anschlussforschungen zugegriffen werden kann. Forschende der digitalen Geisteswissenschaften sehen sich nicht dazu in der Lage, zugleich an urheberrechtlich geschützten Werken zu forschen und die Grundsätze der guten wissenschaftlichen Praxis zu beachten.¹⁵⁰ Im Folgenden wird untersucht, welche Möglichkeiten bestehen bzw. welche Grenze zu beachten sind. Die transparente Aufbewahrung der Daten für einen angemessenen Zeitraum wird in Abschnitt 5.1., die Überprüfbarkeit der wissenschaftlichen Forschung anhand der Korpora zu einem späteren Zeitpunkt in Abschnitt 5.2. und die Herausgabe der Forschungsdaten zu Zwecken von Anschlussforschungen in Abschnitt 5.3. analysiert.

5.1 Aufbewahrung der Korpora

Vor der Umsetzung der DSM-RL ins nationale Urheberrecht mussten die Korpora nach Abschluss der Forschungsarbeiten entweder gelöscht oder an Einrichtungen nach §§ 60e, f UrhG (u. a. Bibliotheken, Museen und Archive) zur Aufbewahrung übermittelt werden. Insofern konnten Forschende die

¹⁵⁰ So auch Raue/Schöch, RuZ 2020, S. 118, 120 ff.

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 41–55. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.e>. License: CC BY 4.0

Korpora zwar nicht selbst behalten, sie durften aber langfristig aufbewahrt werden, wie es die gute wissenschaftliche Praxis erfordert. Im Zuge der Umsetzung der DSM-RL wurde diese als solche bezeichnete Löschungspflicht¹⁵¹ jedenfalls bzgl. nicht-kommerziell tätiger Wissenschaftsinstitutionen und Kulturerbe-Einrichtungen (§ 60d Abs. 2 S. 1, Abs. 3 Nr. 1 UrhG n. F.) abgeschafft. Es ist diesen Personenkreisen bzw. Einrichtungen nun gestattet, die Korpora aufzubewahren, „solange sie für Zwecke der wissenschaftlichen Forschung oder zur Überprüfung wissenschaftlicher Erkenntnisse erforderlich sind“, § 60d Abs. 5 UrhG.¹⁵² Diese Anforderlichkeit erfasst in der Regel den von der guten wissenschaftlichen Praxis geforderten Zeitrahmen von zehn Jahren, unterliegt aber der Einschätzungsprärogative der Forschenden.¹⁵³ Das kann etwa dadurch bewerkstelligt werden, dass die standardmäßige Aufbewahrungsdauer zehn Jahre beträgt und Abweichungen, sei es eine Verlängerung oder Verkürzung der Aufbewahrungsfrist, durch die Forschenden gesondert begründet werden

Kurz zusammengefasst...

	§ 60d Abs. 3 UrhG a. F. (2018)	§ 60d Abs. 5 UrhG n. F. (2021)
	Übermittlung an bestimmte Institutionen	Eigene Aufbewahrung durch Forschungs- und Kulturerbe-Einrichtungen Fremdarchivierung bei Projektkooperationen
		Übermittlung zu Aufbewahrungszwecken
	Eigene Aufbewahrung	

¹⁵¹ In Bezug auf den Namen sind deswegen Bedenken zu äußern, weil die Korpora in jedem Fall zur dauerhaften Aufbewahrung übermittelt werden durfte und dann nicht gelöscht werden mussten.

¹⁵² Im nicht-privilegierten Bereich sind die Korpora weiterhin zu löschen, § 44b Abs. 2 S. 2 UrhG.

¹⁵³ Raue, ZUM 2021, S. 793, 798 f.; ähnlich Wirth, ZUM 2020, S. 585, 590 in Bezug auf den gleichlautenden Stand des ersten Diskussionsentwurfs.

müsste.¹⁵⁴ Ob nach der Gesetzesreform weiterhin eine Fremdarchivierung möglich ist, wie sie zur Einhaltung der guten wissenschaftlichen Praxis erforderlich ist,¹⁵⁵ ist nicht eindeutig. Die Fremdarchivierung ist – der Gesetzesbegründung zufolge – wohl intendiert, aber nicht ausdrücklich erlaubt.¹⁵⁶ Möglich ist allerdings, die Korpora bereits während der Forschungsarbeiten auf Servern von Wissenschaftseinrichtungen zu speichern und sie dann dort zu belassen.¹⁵⁷ Die Einrichtungen können sich wiederum auf die Erlaubnis berufen, die Korpora für einen angemessenen Zeitraum aufzubewahren (§ 60d Abs. 5 UrhG).

5.2 Überprüfbarkeit der wissenschaftlichen Forschung

Die Überprüfbarkeit ist wesentlicher Bestandteil der wissenschaftlichen Forschung und sichert ihre Qualität und Validität. Aus urheberrechtlicher Perspektive gehen mit der Zugänglichmachung zu Überprüfungszwecken allerdings in vielen Fällen urheberrechtliche Verwertungshandlungen wie Vervielfältigungen, § 16 Abs. 1 UrhG, oder öffentliche Zugänglichmachungen, § 19a UrhG, einher.

§ 60d Abs. 4 UrhG enthält zwar explizit die Überprüfbarkeit, sie bezieht sich aber auf Grundlage ihrer systematischen Stellung vor Abs. 5, der die Speicherung regelt, nur auf Peer-Review-Verfahren, die entweder vor Abschluss der Forschungsarbeiten oder unmittelbar im Anschluss daran stattfinden.¹⁵⁸ Die Gesetzesbegründung erwähnt dazu ausdrücklich, dass die öffentliche Zugänglichmachung nach Abschluss der Forschungsarbeiten zu beenden ist.¹⁵⁹ Die Überprüfung wissenschaftlicher Forschung muss aber zu einem späteren Zeitpunkt möglich sein, denn sie wird sogar unter die Freiheit von Wissenschaft und Forschung gefasst, Art. 5 Abs. 3 S. 1 GG.¹⁶⁰ Es erschiene auch widersprüchlich, wenn der Gesetzgeber einerseits eine Aufbewahrungsmöglichkeit vorsähe, die gerade den Zweck einer Überprüfbarkeit und Referenzierbarkeit hat,¹⁶¹ die Forschungsdaten aber andererseits nicht zur Überprüfung wissenschaftlicher Forschung verwendet werden dürften.¹⁶² Deswegen muss § 60d Abs. 4 UrhG so ausgelegt werden, dass sie die Überprüfung und die damit

¹⁵⁴ Vgl. bereits Andresen et al., ZfdG im Erscheinen 2022.

¹⁵⁵ Vgl. bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 197 f.

¹⁵⁶ Vgl. bereits Kleinkopf/Pflüger, ZUM 2021, S. 643, 646 f.; BT-Drs. 19/27426, S. 97.

¹⁵⁷ Vgl. bereits Andresen et al., ZfdG im Erscheinen 2022.

¹⁵⁸ S. bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 198.

¹⁵⁹ BT-Drs. 19/27426, S. 97.

¹⁶⁰ So schon Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 198; u. a. Gärditz in Maunz/Dürig, GG, Art. 5 Abs. 3 Rn. 94 mwN.

¹⁶¹ BT-Drs. 18/12329, S. 41.

¹⁶² So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 198; ähnlich Raue, CR 2017, S. 656, 661.

einhergehenden urheberrechtlich relevanten Handlungen unabhängig davon erlaubt, ob die Überprüfung unmittelbar nach Abschluss der Forschungsarbeiten oder zu einem späteren Zeitpunkt erfolgt,¹⁶³ und die Zugänglichmachung erst mit Abschluss der Überprüfung, nicht aber mit Abschluss der Forschungsarbeiten, zu beenden ist.

5.3 Nachnutzung von TDM-Korpora

Negative Auswirkungen auf die Forschung mit TDM hat insbesondere die Unsicherheit bzgl. der Frage danach, ob der Rechtsrahmen in Bezug auf eine Weiterverwendung der Korpora den wissenschaftlichen Ansprüchen genügt:¹⁶⁴ Gerade die Verfügbarkeit von Forschungsdaten für Anschlussnutzungen ist ein nicht wegzudenkender Teil eines Forschungsprozesses, der zur guten wissenschaftlichen Praxis zählt¹⁶⁵ und von den FAIR-Prinzipien verlangt wird.¹⁶⁶ Diese besagen, dass die Daten „*Findable*“, „*Accessible*“, „*Interoperable*“ sowie „*Reusable*“, also auffindbar, zugänglich, interoperabel und wiederverwendbar, sein müssen, um eine effiziente Wissensweitergabe zu ermöglichen.¹⁶⁷ Als wichtiger Bestandteil von Open Access¹⁶⁸ wird der Datenzugang und damit die Weiterverwendbarkeit auch regelmäßig zur Förderbedingung von Forschungsprojekten bestimmt.¹⁶⁹ Die Nachnutzbarkeit ist aber auch aus ökonomischen Gründen wünschenswert: Die Datenaufbereitung nimmt einen wesentlichen Teil der Forschungsarbeiten ein und ist demnach besonders kostenintensiv.¹⁷⁰ Um einerseits einen möglichst weiten Nutzen aus investierten Geldern

¹⁶³ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 198.

¹⁶⁴ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 196 f.; Gärtner/Kleinkopf/Andresen/Hermann, Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-9) 2021, S. 10 ff., 12.; allgemein zu negativen Auswirkungen der restriktiven Zulassung von TDM-Forschung Handke/Guibault/Vallbb in Schmidt/Dobrevá (Hrsg.), New avenues for electronic publishing in the age of infinite collections and citizen science, 127 ff.; Haucap/Loebert/Spindler/Thorwarth, Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, S. 45 f.; Raue, GRUR 2017, S. 11, 13.

¹⁶⁵ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199; Deutsche Forschungsgemeinschaft, Leitlinien zur Sicherung guter wissenschaftlicher Praxis (2019), S. 18 f. Leitlinie 18; de la Durantaye/Raue, RuZ 2020, S. 83, 90.

¹⁶⁶ Wilkinson et al., Scientific data 2016, S. 160018.

¹⁶⁷ Ebd.

¹⁶⁸ Euler, RuZ 2020, S. 56, 56 ff.

¹⁶⁹ Dazu bereits S. 6 f.

¹⁷⁰ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

zu ziehen und andererseits auch inhaltlich von bereits betriebener Forschung zu profitieren, ist es daher sowohl wissenschaftlich notwendig als auch unter ökonomischen Gesichtspunkten effizient, Korpora – jedenfalls anteilig – nachnutzbar zu machen.¹⁷¹

Die Nachnutzbarkeit von TDM-Korpora ist gesetzlich nicht ausdrücklich vorgesehen. Die Möglichkeit der Nachnutzung ist auch in der juristischen Literatur unklar, obwohl sie mehrfach gefordert oder zumindest für wichtig befunden wurde.¹⁷² Bei der der Beantwortung dieser Rechtsfrage ist zu beachten, dass seitens der Urheber und anderen Rechteinhabern ein berechtigtes Interesse daran besteht, dass keine Parallel-Datenbanken ihrer Schutzgegenstände entstehen.¹⁷³ Dass die Interessen dieser Rechteinhaber nicht über Gebühr beeinträchtigt werden dürfen, gibt auch der urheberrechtliche Dreistufentest, der in Art. 5 Abs. 5 InfoSoc-RL kodifiziert wurde, vor.¹⁷⁴ Dieser gilt auch im Bereich des TDM, Art. 7 Abs. 2 DSM-RL i. V. m. Art. 24, 25 DSM-RL. Zur Verfügbarmachung der Korpora an Dritte, die ebenfalls Forschungszwecke verfolgen, bedarf es einer urheberrechtlichen Schranke, denn eine Übermittlung von Auszügen der Korpora beinhaltet Vervielfältigungen (§ 16 UrhG) und, je nach Empfängerkreis, öffentliche Zugänglichmachungen (§ 19a UrhG) geschützter Anteile jener Werke, die im Korpus enthalten sind. Einige der vorhandenen urheberrechtlichen Schrankenbestimmungen können zur Ermöglichung einer Nachnutzung von TDM-Korpora in Erwägung gezogen werden. Dazu zählen § 60e Abs. 4 UrhG (der in Bibliotheken sog. Terminal-Kopien ermöglicht) sowie die Schranke des § 60c UrhG (die Wissenschaftsschranke im deutschen UrhG).

5.3.1 Terminal-Kopien, § 60e Abs. 4 UrhG

Es wurde erwogen, die sog. Terminal-Schranke, § 60e Abs. 4 UrhG, heranzuziehen.¹⁷⁵ Diese Vorschrift erlaubt Bibliotheken, ihren Nutzenden an Terminals in ihren Räumlichkeiten Bestandswerke zugänglich zu machen (S. 1) und

¹⁷¹ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199. Zustimmung auch Raue, CR 2017, S. 656, 659; der aufgrund der investierten finanziellen Ressourcen ein öffentliches Interesse an der Weiterverwendung begründet und eine Löschungspflicht als Verschwendung von Ressourcen bezeichnet, vgl. Raue, GRUR 2017, S. 11, 15.

¹⁷² Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, § 60d Rn. 13; Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 18; de la Durantaye/Raue, RuZ 2020, S. 83, 94; Hauk/Pflüger, ZUM 2020, S. 383, 385 Raue, CR 2017, S. 656, 660 f; Spindler, ZGE 2018, S. 273, 285;

¹⁷³ Raue, CR 2017, S. 656, 659; BT-Drs. 18/12329, S. 41.

¹⁷⁴ Zu dem Dreistufentest vgl. ausführlicher S. 38 ff.

¹⁷⁵ Spindler, ZGE 2018, S. 273, 285; Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60d Rn. 18; Raue, CR 2017, S. 656, 661.

Nutzenden der Terminals, Anschlusskopien von bis zu zehn Prozent je Sitzung anzufertigen (S. 2). Die Anwendung dieser Vorschrift hätte auf der einen Seite den großen Vorteil, dass Forschende vor Ort anhand der vollständigen Korpora interessengerecht diejenigen Anteile auswählen können, die für ihre Interessen geeignet sind. Auf der anderen Seite ist die Norm an die Nutzung vor Ort („*Closed-Room-Zugang*“)¹⁷⁶ gebunden. Das ist einerseits in digitalen Kontexten und andererseits in einer internationalen Forschungslandschaft nicht zumutbar.¹⁷⁷

5.3.2 Nicht-kommerzielle, wissenschaftliche Forschung, § 60c UrhG

Die Schranke zugunsten der nicht-kommerziellen wissenschaftlichen Forschung, § 60c UrhG, ermöglicht es, bis zu 15 Prozent von Werken bzw. vollständige Werke geringen Umfangs zu vervielfältigen und einem bestimmten, abgegrenzten Personenkreis für deren eigene wissenschaftliche Forschung öffentlich zugänglich zu machen. Im Gegensatz zu § 50d Abs. 4 UrhG sieht § 60c UrhG also keine Ortsbindung vor, sondern ermöglicht auch die digitale Bereitstellung von urheberrechtlichen Schutzgegenständen. Im Folgenden wird dargelegt, wieso § 60c UrhG in Anschluss an Nutzungshandlungen nach §§ 60d, 44b UrhG Anwendung finden kann.

5.3.2.1 Anwendbarkeit

§ 60c UrhG wurde in Umsetzung von Art. 5 Abs. 3 lit. a InfoSoc-RL verabschiedet. Die auf dieser InfoSoc-RL basierenden Schranken sind trotz der jüngeren DSM-RL anwendbar: Auf Grundlage des Art. 24, 25 DSM-RL können auf Grundlage der InfoSoc- und Datenbank-RL umfassendere Regelungen zu den durch die DSM-RL vorgesehenen Schranken geschaffen werden. Konkret bedeutet dies, dass das deutsche Recht über die DSM-RL hinaus Regelungen treffen darf, sofern die Anforderungen der InfoSoc-RL eingehalten werden.

5.3.2.2 Gesetzgeberische und inhaltliche Intention

Die Intention des europäischen Gesetzgebers spricht dafür, die Schranke zugunsten wissenschaftlicher Forschung für eine Nachnutzbarkeit heranzuziehen. Die DSM-RL besagt dazu: *„Die Nutzung zum Zwecke der wissenschaftlichen Forschung außerhalb des Text und Data Mining, etwa die Begutachtung unter wissenschaftlichen Fachkollegen und gemeinsame Forschungsarbeiten, sollte nach wie vor unter die Ausnahme oder Beschränkung im Sinne von Artikel 5 Absatz 3 Buchstabe a der Richtlinie 2001/29/EG fallen, sofern diese Bestimmung*

¹⁷⁶ Raue/Schöch, RuZ 2020, S. 118, 122.

¹⁷⁷ Ähnlich auch Raue/Schöch, RuZ 2020, S. 118, 122.

anwendbar ist“.¹⁷⁸ Der EU-Gesetzgeber hatte Nutzungen außerhalb des TDM selbst also nicht nur intendiert,¹⁷⁹ er ging darüber hinaus sogar davon aus, dass Mitgliedsstaaten Nutzungen, die über die Vorgaben der DSM-RL hinausgehen, auf Grundlage der InfoSoc-RL ermöglichen können. Zudem erwähnt § 60d Abs. 5 UrhG ausdrücklich, dass die Korpora aufbewahrt werden dürfen, wenn das für Zwecke der wissenschaftlichen Forschung erforderlich ist. In der Gesetzesbegründung ist außerdem formuliert: „Die Nutzung zum Zwecke der wissenschaftlichen Forschung außerhalb des Text und Data Mining, etwa für die Begutachtung unter wissenschaftlichen Fachkollegen und für gemeinsame Forschungsarbeiten, fällt nach wie vor unter die Ausnahme des § 60c UrhG, sofern die Voraussetzungen dieser Vorschrift erfüllt sind (vergleiche ErwG 15 Satz 5 DSM-RL)“.¹⁸⁰ Darauf könnte insofern auch die Herausgabe von Korpusauszügen gestützt werden, die eine wissenschaftliche Zwecksetzung hat, schließlich ist die Aufzählung in der Gesetzesbegründung nicht abschließend. Systematisch und inhaltlich ist § 60c UrhG damit die passende Möglichkeit, TDM-Textbestände nachnutzbar zu machen.

5.3.2.3 Zulässigkeit der Kumulation urheberrechtlicher Schranken

Als problematisch könnte allerdings erachtet werden, dass eine Ausweitung von Schrankenbestimmungen in einem Urheberrecht erfolgen soll, das tendenziell die Rechteinhaber fokussiert. Doch die Kumulation des § 60d Abs. 5 UrhG mit § 60c UrhG ist denkbar: In den letzten Jahren ist eine Entwicklung dahingehend erkennbar, die Nutzungsinteressen als legitim bewertet und rechtlich fördert.¹⁸¹ Auf ein hohes Schutzniveau zielte noch die InfoSoc-RL ab, die mittlerweile bereits 20 Jahre alt ist.¹⁸² Demgegenüber gestaltet Art. 3 DSM-RL die Erlaubnis des TDM im Urheberrecht der EU-Mitgliedsstaaten nicht nur – wie davor üblich – fakultativ, sondern verpflichtend aus.¹⁸³ Darüber hinaus entfernt sich das Urheberrecht langsam aber sicher von der traditionell engen Schrankenauslegung:¹⁸⁴ Mittlerweile gehen nicht nur Stimmen der Literatur von einer

¹⁷⁸ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 198 f.; DSM-RL Erw. 15 S. 5.

¹⁷⁹ So auch Raue, ZUM 2019, S. 684, 688; Raue, CR 2017, S. 656, 659 ff.

¹⁸⁰ BT-Drs. 19/27426, S. 97.

¹⁸¹ So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

¹⁸² Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

¹⁸³ Ebd.

¹⁸⁴ Zur früheren engen Schrankenauslegung u. a. Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, Vor §§ 44a ff. Rn. 36; BGH, Urt. v. 03.04.1968 – I ZR 83/66, Kandinsky = GRUR 1968, 607, S. 608 f.; BGH, Landesversicherungsanstalt = GRUR 1972, 614, S. 616; BGH, Urt. v. 18.04.1985 – I ZR 24/83, Schulfunksendung = GRUR 1985, 874, S. 875; BGH, Urt. v. 06.06.1991 – I ZR 26/90, Liedersammlung = GRUR 1991, 903, S. 905; BGH,

Schrankenlegung nach ihrem Sinn und Zweck aus.¹⁸⁵ Auch der EuGH hat sich dementsprechend positioniert.¹⁸⁶

Dem dargestellten Ansatz steht auch nicht der bereits erwähnte und in Art. 5 Abs. 5 InfoSoc-RL kodifizierte Drei-Stufen-Test entgegen, der sowohl Grundlage für die Normierung urheberrechtlicher Schranken als auch relevant für die richterliche Rechtsauslegung ist.¹⁸⁷ Ihm zufolge müssen urheberrechtliche Schranken auf bestimmte Sonderfälle begrenzt sein (Stufe eins), sie dürfen daneben die normale Verwertung des Werks nicht beeinträchtigen (Stufe zwei) und die berechtigten Interessen des Urhebers nicht unzumutbar verletzen (Stufe drei). Die wissenschaftliche Nachnutzung von TDM-Korpora ist ein derartiger Sonderfall.¹⁸⁸ Es ist auch der Primärmarkt nicht betroffen: Die segmentierten und annotierten Werke eignen sich nicht als Ersatz des ursprünglichen Werks. Dies gilt erst recht, wenn nur Auszüge ausgegeben werden.¹⁸⁹ Die Anforderungen an eine Unzumutbarkeit für die Rechteinhaber, die in Form einer Interessensabwägung geprüft werden,¹⁹⁰ sind tendenziell als hoch anzusehen.¹⁹¹ Berechtigte Interessen an den Korpora bestehen seitens der wissenschaftlichen Forschung und basieren auf der Wissenschaftsfreiheit, Art. 13 GRCh bzw. auf Art. 5 Abs. 3 GG. Demgegenüber sind die Urheber der in den Korpora enthaltenen geschützten Werke zu berücksichtigen, Art. 17 Abs. 2 GRCh bzw. Art. 14 GG. Zugunsten der Nachnutzungsmöglichkeit kann insbesondere die Vergütungspflicht des § 60c UrhG, § 60h UrhG, berücksichtigt

Urt. v. 17.03.1983 – I ZR 186/80, Zoll- und Finanzschulen = GRUR 1983, 562, S. 563; BGH, Urt. v. 30.11.2011 – I ZR 212/10, Blühende Landschaft en = GRUR 2012, 819, S. 822 Rn. 28.

¹⁸⁵ Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, Vor §§ 44a ff. Rn. 37 mwN; tendenziell Dreier in Dreier/Schulze, UrhG, 6. Auflage 2018, Vor § 4a Rn. 7; Stieper, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 65 ff.

¹⁸⁶ EuGH, ECLI:EU:C:2019:623, Funke Medien/Bundesrepublik Deutschland [Afghanistan Papiere] = GRUR 2019, 934, S. 939 69 ff; EuGH, ECLI:EU:C:2019:625, Spiegel Online/Volker Beck [Reformistischer Aufbruch] = GRUR 2019, 940, S. 944 Rn. 53 ff. So bereits Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

¹⁸⁷ Stieper, Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts, S. 73 mwN; BGH, Urt. v. 11.7.2002 – I ZR 255/00, Elektronischer Pressespiegel = GRUR 2002, 963, S. 967 ff.

¹⁸⁸ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

¹⁸⁹ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 199.

¹⁹⁰ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, Vor § 44a Rn. 21; Stieper in Schricker/Loewenheim, UrhG, 6. Auflage 2020, Vor §§ 44a ff. Rn. 30; näher Senfleben, GRUR Int. 2004, S. 200, 206.

¹⁹¹ Senfleben, GRUR Int. 2004, S. 200, 210.

werden, die über die zugehörigen Forschungs- und Kulturerbe-Einrichtungen getragen werden.¹⁹² Das TDM selbst ist zwar entschädigungslos¹⁹³ hinzunehmen. Die Schranke des § 60d UrhG setzt allerdings den rechtmäßigen Zugang zu geschützten Ausgangsmaterialien voraus. Durch den Erwerb des rechtmäßigen Zugangs fließen den Rechteinhabern mittelbare Einkünfte zu. Bei einer Nachnutzung soll dieser Zugang gerade erst gewährt werden, weswegen der (oft vergütungspflichtig erworbene) rechtmäßige Zugang, der nach gesetzgeberischer Intention im ersten Schritt erfolgt, ausbleibt. Über diese Divergenz hilft die Vergütungspflicht des § 60c, § 60h UrhG, hinweg. Aufgrund dieser Erwägungen ist eine Nachnutzung den Rechteinhabern zuzumuten. Insofern entspricht die Zugänglichmachung von Korpusauszügen den Anforderungen des Drei-Stufen-Tests.

Über das Besagte hinaus wurde die Möglichkeit, Schranken miteinander zu kombinieren, wie es bei der Anwendung des § 60c UrhG auf § 60d UrhG der Fall wäre, auch schon in der Vergangenheit durch den EuGH thematisiert und ist nach dessen Ansicht möglich, wenn die Anforderung der jeweiligen Bestimmung erfüllt sind.¹⁹⁴ Die Kumulation zweier Schranken ist von der DSM-RL in Erw. 15 auch explizit bezweckt.

5.3.3 Auszugskonzept

Im Rahmen von XSample wurde ein Auszugskonzept erstellt, das in das Datenrepositorium der Universität Stuttgart integriert werden kann. Eine besondere Rolle nehmen bei dem vorgestellten Konzept diejenigen Einrichtungen an, die die Korpora für die Forschenden aufbewahren, z. B. wissenschaftliche Bibliotheken. Das Konzept beinhaltet, dass die Einrichtung, die die Korpora aufbewahrt, Forschenden die Möglichkeit bietet, Korpusauszüge zu Zwecken der wissenschaftlichen Nachnutzung zu erhalten. Eine auszugsweise Übermittlung der Korpora ist zwar nicht ausreichend, um vollständige Studien daran vornehmen zu können, doch kann sie Anhaltspunkte dafür bieten, ob sich Forschungsmaterial für bestimmte Forschungen eignet. Im ersten Anwendungsfall wäre diese Vorgehensweise etwa dafür nützlich, um festzustellen, ob relevante textuelle Phänomene in den betreffenden Texten enthalten sind oder ob Art und Qualität der Annotationsdaten für die eigene Forschung geeignet sind.¹⁹⁵ Über ein Such-Tool in einem Web-Interface kann präzise ausgedrückt

¹⁹² Stieper in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, Vor §§ 44a ff. Rn. 31; Senftleben, GRUR Int. 2004, S. 200, 211; EuGH, ECLI:EU:C:2014:2196, Eugen Ulmer = GRUR 2014, 1078, S. 1080 Rn. 48.

¹⁹³ Anders verhält es sich im alten Recht, § 60h Abs. 1 UrhG a. F.

¹⁹⁴ Ähnlich EuGH, ECLI:EU:C:2014:2196, Eugen Ulmer = GRUR 2014, 1078.

¹⁹⁵ Kleinkopf/Jacke/Gärtner, MMR 2021, S. 196, 200.

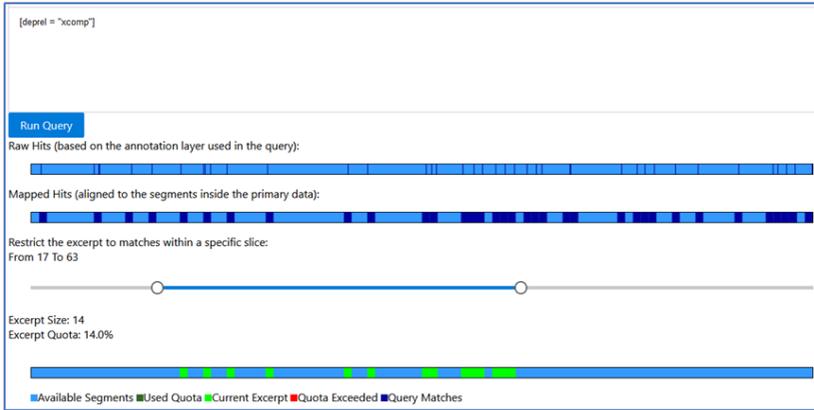


Abbildung 3: Auszugskonzept. Quelle: Eigene Darstellung von Markus Gärtner, aus: Gärtner et al., Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-9) 2021, S. 17; CC BY 4.0 International.

werden, welche Passagen für das jeweilige Forschungsinteresse relevant sind (Abbildung 3). Das wird auf Grundlage der Annotationsdaten bestimmt. Daraufhin kann der Forschende innerhalb der prozentualen Grenzen die Passagen auswählen, die in seinem Korpusauszug enthalten sein sollen. Diese werden dann zum Download bereitgestellt.¹⁹⁶ Mitgeliefert wird dabei eine Text-Datei, die die Quellenangaben der in den Auszügen enthaltenen Werke bereitstellt.

5.3.3.1 Rechtliche Eignung des Auszugskonzepts

Im Folgenden wird geprüft, ob die Norm alle Rechtsverletzungen, die mit der Herausgabe von Korpusauszügen einhergehen, adäquat abdeckt. Relevant werden bei der Herausgabe von Korpusauszügen die Rechte der Urheber oder Inhaber verwandter Schutzrechte der in den Korpora enthaltenen Schutzgegenstände. Im dargestellten Ansatz werden die Rechte an den Korpora selbst nicht berücksichtigt, d. h. der Ansatz basiert auf der Prämisse, dass Forschende, die dazu bereit sind, Auszüge herauszugeben, nicht auf die Rechte an ihren Korpora bestehen.

¹⁹⁶ Ausführlich dazu Gärtner/Kleinkopf/Andresen/Hermann, Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-9) 2021, S. 10 ff., 14 ff; Andresen et al., ZfdG im Erscheinen 2022.

5.3.3.2 Betroffene Verwertungshandlungen

5.3.3.2.1 *Vervielfältigung, § 16 UrhG, und Bearbeitung, § 23 UrhG*

Eine Übermittlung von Schutzgegenständen beinhaltet bei digitaler Übermittlung stets eine Vervielfältigung, § 16 Abs. 1 UrhG. Wie deutlich wurde, handelt es sich bei den in den Korpora enthaltenen Forschungsdaten in manchen Fällen um Bearbeitungen der Ursprungswerke. Diese dürfen nach dem Gesetz nur mit Zustimmung der Rechteinhaber verwertet werden. Eine solche Verwertung findet allerdings statt, wenn Auszüge vervielfältigt und öffentlich zugänglich gemacht werden. In diesen Fällen ist § 60c UrhG, wie bereits in Bezug auf §§ 60d, 44b UrhG argumentiert wurde, im Wege einer extensiven, europarechtskonformen Auslegung auch auf Bearbeitungen zu beziehen. Das betrifft also Werkexemplare, die das Werk nicht vollkommen originalgetreu wiedergeben, denn das Unionsrecht macht keine Unterschiede zwischen identischen und abgewandelten Kopien.¹⁹⁷ Deswegen ist § 60c UrhG so zu verstehen, dass er auch die Vervielfältigung von Bearbeitungen erfasst.

5.3.3.3.2 *Öffentliche Zugänglichmachung, § 19a UrhG*

In einer Implementierung des vorgestellten Auszugskonzeptes, bei der Auszüge automatisiert generiert und herausgegeben werden, werden die erzeugten Auszüge dadurch öffentlich zugänglich gemacht, § 19a UrhG, indem sie zum anteiligen Download bereitgestellt werden. Das ist auch der Fall, wenn einzelne Personen erst zeitlich nach Prüfung der Zweckbindung die Möglichkeit zum Download eines generierten Auszugs erhalten. Das liegt daran, dass es genügt, dass dieselben Schutzgegenstände nur Mitgliedern der Öffentlichkeit zur gleichen Zeit zugänglich sind.¹⁹⁸ § 60c UrhG rechtfertigt derartige öffentliche Zugänglichmachungen. Die Norm muss allerdings ebenso extensiv ausgelegt werden, sodass auch Bearbeitungen (§ 23 Abs. 1 UrhG), also nicht nur identische Vervielfältigungen (§ 16 Abs. 1 UrhG) öffentlich zugänglich gemacht werden dürfen.

5.3.3.3.3 *Quellenangabe, § 63 UrhG*

Das Namensnennungsrecht, § 63 Abs. 1 S. 1 UrhG, gilt – im Gegensatz zu § 60d UrhG – im Bereich des § 60c UrhG. Deswegen muss in den Auszügen

¹⁹⁷ Vgl. dazu bereits S. 16 f.

¹⁹⁸ v. Ungern-Sternberg in Schrickler/Loewenheim, UrhG, 6. Auflage 2020, § 19a Rn. 69 f., § 15 Rn. 69; EuGH, ECLI:EU:C:2014:2196, Eugen Ulmer = GRUR 2014, 1078, S. 1080 Rn. 41; EuGH, ECLI:EU:C:2014:76, Nils Svensson ua/Retriever Sverige = GRUR 2014, 360, S. 361 Rn. 19.

grundsätzlich darauf hingewiesen werden, wer Urheber der auszugsweise enthaltenen Werke ist und welche Veränderungen an den Werken erfolgt sind. Weil die Vervielfältigungen, d. h. die Korpora, die Quelle schon nicht enthalten müssen, könnte zwar angenommen werden, dass auch die Quellenangabe in Auszügen der erstellten Korpora unmöglich wäre und die Pflicht zur Quellenangabe aus diesem Grund entfalle.¹⁹⁹ Von XSample wird dennoch die Anfügung einer „legal.txt“-Datei vorgeschlagen, die die notwendigen Informationen beinhaltet und den Korpusauszügen beigelegt wird, um etwaigen Risiken vorzubeugen.

5.3.3.2.5 Änderungsverbot, § 62 UrhG

Eine Änderung i. S. d. § 62 UrhG kann auch dann vorliegen, wenn der Sinngehalt des Werks durch die Schrankennutzung verfälscht wird.²⁰⁰ Das kommt bei der Generierung und Herausgabe von Auszügen der Korpora deswegen in Betracht, weil die enthaltenen Schutzgegenstände ggf. aus dem Kontext gerissen werden. Die Interessenabwägung der §§ 62 Abs. 1, 39 Abs. 2 UrhG ergibt allerdings, dass die Interessen der Rechteinhaber nicht beeinträchtigt sind: Das Änderungsverbot soll den Urheber davor schützen, dass sein Werk in der Öffentlichkeit mit Änderungen präsentiert wird, die nicht von ihm stammen, ihm aber zugerechnet werden.²⁰¹ Im Falle der Auszugsgenerierung aus Korpora ist dies nicht zu befürchten: Die Auszüge werden an Forschende herausgegeben, d. h. es ist den Interessenten nicht nur bewusst, sondern auch von ihnen gewünscht, dass es sich dabei um weitreichend bearbeitete Texte handelt. Zudem werden die Auszüge nicht einer breiten Öffentlichkeit, sondern einzelnen Mitgliedern der Öffentlichkeit auf Anfrage zugänglich gemacht. Im Rahmen der Interessenabwägung des § 39 Abs. 2 UrhG ist daher den Nutzungsinteressen der Wissenschaft Vorzug zu gewähren.

5.3.3.2.6 Vergütungspflicht

Die Vervielfältigungen und öffentlichen Zugänglichmachungen, die auf § 60c UrhG beruhen, unterliegen im Gegensatz zu denjenigen, die auf § 60d UrhG gestützt werden, einer Vergütungspflicht, § 60h UrhG, die durch die Einrichtungen zu tragen ist.

¹⁹⁹ Vgl. dazu bereits S. 24.

²⁰⁰ Peukert in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 62 Rn. 9.

²⁰¹ Vgl. bereits oben, Peukert in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 14 Rn. 1.

5.3.3.3 Rechtliche Grenzen des § 60c UrhG

Im letzten Schritt wird dargestellt, wie die Einhaltung der rechtlichen Grenzen technisch sichergestellt wird.

5.3.3.3.1 Empfängerkreis

Die Vervielfältigung und öffentliche Zugänglichmachung für die wissenschaftliche Forschung dürfen nur für einen bestimmt abgegrenzten Personenkreis erfolgen, d. h. Korpora dürfen nicht etwa allgemein in einem Universitätsintranet zum Download bereitgestellt werden.²⁰² Eine Nachnutzung kann insofern nur dergestalt erfolgen, dass der Zugriff zum Auszugs-Auswahl-Tool erst nach individueller Anfrage gewährt wird.

5.3.3.3.2 Bestimmung der Auszüge

Gem. § 60c Abs. 1 UrhG dürfen maximal 15 Prozent eines Werks vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden. Eine Bezugsgröße für die Bemessung des prozentualen Anteils des Auszugs vom Gesamtwerk gibt weder das Gesetz noch die Gesetzesbegründung vor. Es wird daher gemeinhin angenommen, dass die Maßstäbe anzuwenden sind, die bereits für die Vorgängernormen²⁰³ des § 60c UrhG galten.²⁰⁴ Das bedeutet, dass die Seitenzahl des Werks in seinem Gesamtumfang inkl. Einleitung und Vorwort sowie Inhalts-, Literatur- und Sachverzeichnis Bemessungsgrundlage ist.²⁰⁵ Nach BGH-Rechtsprechung sind im Falle von Sprachwerken i. S. d. § 2 Abs. 1 Nr. 1, Abs. 2 UrhG sämtliche Seiten zu berücksichtigen, die keine Leerseiten sind und deren Inhalt überwiegend aus Text besteht.²⁰⁶ Bei mehrbändigen Werken gilt das für jeden Band einzeln.²⁰⁷ Durch Annotationen muss in den Korpora sichergestellt werden, dass die Seitenanzahl der Primärdaten erhalten bleibt, sodass der Umfang von Korpusauszügen berechnet werden kann.

²⁰² Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60c Rn. 9; BT-Drs. 15/837, 34.

²⁰³ Die Vorgängernorm, § 52a Abs. 1 Nr. 1 UrhG, enthielt die Begrenzung auf „*kleine Teile eines Werks*“.

²⁰⁴ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60c Rn. 7.

²⁰⁵ Ebd.; BGH, Urt. v. 28.11.2013 – I ZR 76/12, Meilensteine der Psychologie = GRUR 2014, 549, S. 551 Rn. 24.

²⁰⁶ BGH, Urt. v. 28.11.2013 – I ZR 76/12, Meilensteine der Psychologie = GRUR 2014, 549, S. 551 Rn. 24, 25.

²⁰⁷ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60a Rn. 8; Anton in Spindler/Schuster, Recht der elektronischen Medien, § 60a Rn. 26; Wandtke, NJW 2018, S. 1129, 1134.

5.3.3.3.3 Nutzung vollständiger Werke geringen Umfangs

Gem. § 60c Abs. 3 UrhG gilt die Begrenzung auf 15 Prozent eines Werks nicht für Abbildungen, einzelne Beiträge derselben Fachzeitschrift oder wissenschaftlichen Zeitschrift, sonstige Werke geringen Umfangs und vergriffene Werke. Insofern können auch einzelne Beiträge in Sammelwerken privilegiert werden.²⁰⁸ Bei Druckwerken ist ein Umfang von 25 Seiten noch als gering anzusehen.²⁰⁹ Kürzere Werke dürfen infolgedessen unter gewissen Voraussetzungen vollständig herausgegeben werden. Eine Besonderheit gilt dabei bei Sammelbänden. Denn Kurzgeschichten oder andere Kurzbeiträge, die Bestandteil eines als Gesamtkonzept geplanten Bandes sind und die bereits einzeln betrachtet Werke geringen Umfangs sind, sind daneben auch Bestandteil eines geschützten Sammelwerks. Folglich dürfen Beiträge aus Sammelwerke, die einen Umfang von 25 Seiten nicht übersteigen, nur dann vollständig im Auszug aus dem Korpus enthalten sein, wenn sie darüber hinaus nicht mehr als 15 Prozent des Gesamtumfangs des (geschützten) Sammelwerks ausmachen.²¹⁰ Wenn mehrere Beiträge derselben wissenschaftlichen Zeitschrift verwendet werden sollen, muss weiterhin die 15 Prozent-Regel beachtet werden, wobei das jeweilige Heft als übergeordnete Einheit angesehen werden muss.²¹¹ Das bedeutet, dass auch mehrere Kurzbeiträge eines Gesamtwerks verwendet werden dürfen, sofern diese jeweils unter 25 Seiten betreffen und gleichzeitig insgesamt nicht mehr als 15 Prozent des Gesamtumfangs des Werks ausmachen.

5.3.3.4 Absicherung der Institutionen oder Forschenden bei der Herausgabe der Auszüge

Zuletzt stellt sich die Frage, wie sich die Institutionen, die Korpus-Auszüge herausgeben möchten, rechtlich absichern bzw. was sie beachten müssen. In der Praxis wird die Zweckprüfung in vergleichbaren Fällen dergestalt gehandhabt, dass sich Angehörige von Forschungsorganisationen, die unter die Privilegierungen der nicht-kommerziellen wissenschaftlichen Forschung fallen, über ihre E-Mail-Adresse oder die Zugehörigkeit zu einer Universität oder einer anderen Forschungseinrichtung authentifizieren. Im nächsten Schritt muss auf Anfrage versichert werden, dass die Daten nur für die nicht-kommerzielle wissenschaftliche Forschung verwendet werden. Bereits die Zugehörigkeit zu

²⁰⁸ Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60c Rn. 7.

²⁰⁹ BT-Drs. 18/12329, S. 35, 40.

²¹⁰ So auch von Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60c Rn. 7 eingeräumt; in diese Richtung geht auch die Argumentation von Wandtke, NJW 2018, S. 1129, 1134.

²¹¹ Maßstab ist hingegen nicht ein zusammengefasster Jahresband oder die gesamte Zeitschrift in ihrer gesamten Historie, Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60a Rn. 18.; Schricker in Schricker/Loewenheim, UrhG, 6. Auflage 2020, § 60a Rn. 18.

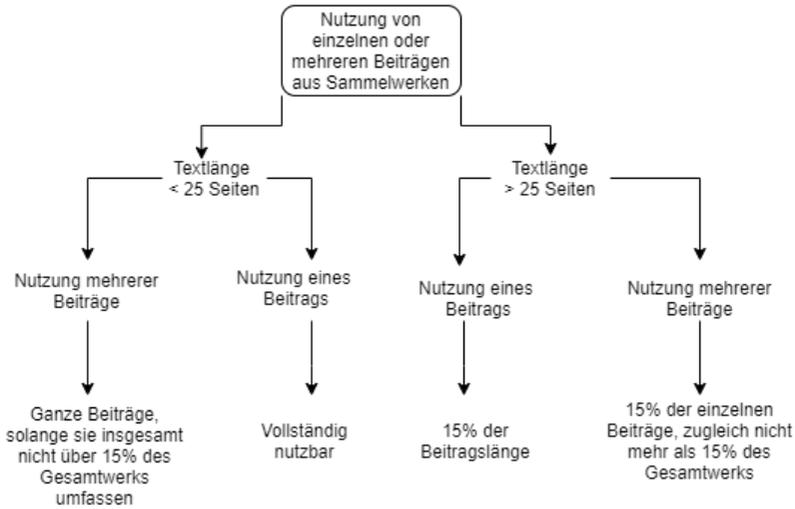


Abbildung 4: Nutzbarkeit kurzer Werke aus Sammelwerken. Quelle: Gärtner et al., Corpus Reusability and Copyright – Challenges and Opportunities, in: Lungen et al. (Hrsg.), Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-9) 2021, S. 17; CC BY 4.0 International.

einer Forschungsorganisation impliziert, dass die Daten im Regelfall auch für die privilegierten Forschungszwecke verwendet werden. Diese Vorgehensweise ist aus praktischen Gründen zu begrüßen.²¹² Anderes dürfte nur für offensichtliche Missbrauchsfälle gelten.²¹³

Kurz zusammengefasst...

Eine Nachnutzbarmachung von Korpusauszügen zu Zwecken der nicht-kommerziellen, wissenschaftlichen Forschung ist auf Grundlage des § 60c UrhG möglich.

²¹² Vgl. dazu auch Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60e Rn. 17, 27 f.; der einräumt, dass sich die Zweckbefolgung nicht kontrollieren lässt und den Bibliotheken keine Prüfpflichten beim Kopienversand auferlegt.

²¹³ So auch Dreier in Dreier/Schulze, UrhG, 7. Auflage 2022, § 60e Rn. 27; in Bezug auf den Kopienversand auf Bestellung gem. § 53a UrhG a. F. Loewenheim in Schrickler/Loewenheim, UrhG, 5. Auflage 2017, § 53a Rn. 9; Jani in Wandtke/Bullinger, UrhG, 4. Aufl. 2014, § 53a Rn. 12. Diese Erwägungen müssen auf die mit dem Kopienversand vergleichbare Situation übertragen werden.

KAPITEL 6

Fazit

Das geltende Urheberrecht lässt sich durch eine extensive Auslegung dergestalt anwenden, dass es den Forschungsprozess des TDM vollständig abdeckt. In einem begrenzten Umfang können auch Korpusauszüge nachgenutzt werden. Gleichwohl sind einige Abschnitte des Forschungsprozesses nur mit ausführlicher rechtlicher Begründung möglich. Insofern bietet das Urheberrecht wohl noch nicht genügend Rechtssicherheit, um eine TDM-Forschung an urheberrechtlichen Schutzgegenständen auch ohne die Einholung von rechtlicher Expertise im Einzelfall zu ermöglichen.

Zitationsvorschlag für dieses Kapitel:

Kleinkopf, F. 2023. *Text- und Data-Mining in den digitalen Geisteswissenschaften: Gutachten zu den urheberrechtlichen Rahmenbedingungen im Rahmen des Forschungsprojektes „XSample – Text- und Data-Mining auf geschützten Werken durch Auszüge transparent erschließen*. Pp. 57. London: Ubiquity Press. DOI: <https://doi.org/10.5334/bcu.f>. License: CC BY 4.0

Anhang: Historie der relevanten Rechtsnormen

Urheberrechts-Wissensgesellschafts-Gesetz, 2018

§ 60c

Wissenschaftliche Forschung

- (1) Zum Zweck der nicht kommerziellen wissenschaftlichen Forschung dürfen bis zu 15 Prozent eines Werkes vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden.
 1. für einen bestimmt abgegrenzten Kreis von Personen für deren eigene wissenschaftliche Forschung sowie.
 2. für einzelne Dritte, soweit dies der Überprüfung der Qualität wissenschaftlicher Forschung dient.
- (2) Für die eigene wissenschaftliche Forschung dürfen bis zu 75 Prozent eines Werkes vervielfältigt werden.
- (3) Abbildungen, einzelne Beiträge aus derselben Fachzeitschrift oder wissenschaftlichen Zeitschrift, sonstige Werke geringen Umfangs und vergriffene Werke dürfen abweichend von den Absätzen 1 und 2 vollständig genutzt werden.

- (4) Nicht nach den Absätzen 1 bis 3 erlaubt ist es, während öffentlicher Vorträge, Aufführungen oder Vorführungen eines Werkes diese auf Bild- oder Tonträger aufzunehmen und später öffentlich zugänglich zu machen.

§ 60d

Text und Data Mining

- (1) Um eine Vielzahl von Werken (Ursprungsmaterial) für die wissenschaftliche Forschung automatisiert auszuwerten, ist es zulässig:
 1. das Ursprungsmaterial auch automatisiert und systematisch zu vervielfältigen, um daraus insbesondere durch Normalisierung, Strukturierung und Kategorisierung ein auszuwertendes Korpus zu erstellen, und
 2. das Korpus einem bestimmt abgegrenzten Kreis von Personen für die gemeinsame wissenschaftliche Forschung sowie einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung öffentlich zugänglich zu machen. Der Nutzer darf hierbei nur nichtkommerzielle Zwecke verfolgen.
- (2) Werden Datenbankwerke nach Maßgabe des Absatzes 1 genutzt, so gilt dies als übliche Benutzung nach § 55a Satz 1. Werden unwesentliche Teile von Datenbanken nach Maßgabe des Absatzes 1 genutzt, so gilt dies mit der normalen Auswertung der Datenbank sowie mit den berechtigten Interessen des Datenbankherstellers im Sinne von § 87b Absatz 1 Satz 2 und § 87e als vereinbar.
- (3) Das Korpus und die Vervielfältigungen des Ursprungsmaterials sind nach Abschluss der Forschungsarbeiten zu löschen; die öffentliche Zugänglichmachung ist zu beenden. Zulässig ist es jedoch, das Korpus und die Vervielfältigungen des Ursprungsmaterials den in den §§ 60e und 60f genannten Institutionen zur dauerhaften Aufbewahrung zu übermitteln.

Richtlinie zur Anpassung des Urheberrechts und der verwandten Schutzrechte an die Erfordernisse des digitalen Binnenmarkts, 2019

Artikel 3

Text und Data Mining zum Zwecke der wissenschaftlichen Forschung:

- (1) Die Mitgliedstaaten sehen eine Ausnahme von den in Artikel 5 Buchstabe a und Artikel 7 Absatz 1 der Richtlinie 96/9/EG, Artikel 2 der Richtlinie 2001/29/EG, und Artikel 15 Absatz 1 der vorliegenden Richtlinie festgelegten Rechten für Vervielfältigungen und Entnahmen vor, die durch Forschungsorganisationen und Einrichtungen des Kulturerbes von Werken oder sonstigen Schutzgegenständen, zu denen sie rechtmäßig

Zugang haben, zum Zwecke der wissenschaftlichen Forschung für die Text und Data Mining vorgenommen werden.

- (2) Vervielfältigungen und Entnahmen von Werken oder sonstigen Schutzgegenständen, die gemäß Absatz 1 angefertigt wurden, sind mit angemessenen Sicherheitsvorkehrungen zu speichern und dürfen zum Zwecke der wissenschaftlichen Forschung, auch zur Überprüfung wissenschaftlicher Erkenntnisse, aufbewahrt werden.
- (3) Die Rechteinhaber müssen Maßnahmen durchführen können, um die Sicherheit und Integrität der Netze und Datenbanken zu wahren, in denen die Werke oder sonstigen Schutzgegenstände gespeichert sind. Diese Maßnahmen dürfen über das für die Verwirklichung dieses Ziels Notwendige nicht hinausgehen.
- (4) Die Mitgliedstaaten wirken darauf hin, dass Rechteinhaber, Forschungsorganisationen und Einrichtungen des Kulturerbes einvernehmlich bewährte Vorgehensweisen bei der Umsetzung der in Absatz 2 genannten Verpflichtung bzw. die Durchführung der in Absatz 3 genannten Maßnahmen definieren.

Artikel 4

Ausnahmen und Beschränkungen für das Text und Data Mining:

- (1) Für zum Zwecke des Text und Data Mining vorgenommene Vervielfältigungen und Entnahmen von rechtmäßig zugänglichen Werken und sonstigen Schutzgegenständen sehen die Mitgliedstaaten eine Ausnahme oder Beschränkung von den Rechten vor, die in Artikel 5 Buchstabe a und Artikel 7 Absatz 1 der Richtlinie 96/9/EG, Artikel 2 der Richtlinie 2001/29/EG, Artikel 4 Absatz 1 Buchstaben a und b der Richtlinie 2009/24/EG und Artikel 15 Absatz 1 der vorliegenden Richtlinie niedergelegt sind.
- (2) Vervielfältigungen und Entnahmen nach Absatz 1 dürfen so lange aufbewahrt werden, wie es für die Zwecke des Text und Data Mining notwendig ist.
- (3) Die Ausnahmen und Beschränkungen nach Absatz 1 finden Anwendung, sofern die jeweiligen Rechteinhaber die in Absatz 1 genannten Werke und sonstigen Schutzgegenstände nicht ausdrücklich in angemessener Weise, etwa mit maschinenlesbaren Mitteln im Fall von online veröffentlichten Inhalten, mit einem Nutzungsvorbehalt versehen haben.
- (4) Dieser Artikel lässt die Anwendung von Artikel 3 unberührt.

Gesetz zum Urheberrecht im digitalen Binnenmarkt, 2021

§ 44b Text und Data Mining:

- (1) Text und Data Mining ist die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus

Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.

- (2) Zulässig sind Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Die Vervielfältigungen sind zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind.
- (3) Nutzungen nach Absatz 2 Satz 1 sind nur zulässig, wenn der Rechteinhaber sich diese nicht vorbehalten hat. Ein Nutzungsvorbehalt bei online zugänglichen Werken ist nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt.

§ 60d Text und Data Mining für Zwecke der wissenschaftlichen Forschung:

- (1) Vervielfältigungen für Text und Data Mining (§ 44b Absatz 1 und 2 Satz 1) sind für Zwecke der wissenschaftlichen Forschung nach Maßgabe der nachfolgenden Bestimmungen zulässig.
- (2) Zu Vervielfältigungen berechtigt sind Forschungsorganisationen. Forschungsorganisationen sind Hochschulen, Forschungsinstitute oder sonstige Einrichtungen, die wissenschaftliche Forschung betreiben, sofern sie:
 1. nicht kommerzielle Zwecke verfolgen.
 2. sämtliche Gewinne in die wissenschaftliche Forschung reinvestieren oder
 3. im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind.
- (3) Nicht nach Satz 1 berechtigt sind Forschungsorganisationen, die mit einem privaten Unternehmen zusammenarbeiten, das einen bestimmenden Einfluss auf die Forschungsorganisation und einen bevorzugten Zugang zu den Ergebnissen der wissenschaftlichen Forschung hat.
- (4) Zu Vervielfältigungen berechtigt sind ferner:
 1. Bibliotheken und Museen, sofern sie öffentlich zugänglich sind, sowie Archive und Einrichtungen im Bereich des Film- oder Tonerbes (Kulturerbe-Einrichtungen),
 2. einzelne Forscher, sofern sie nicht kommerzielle Zwecke verfolgen.
- (5) Berechtigte nach den Absätzen 2 und 3, die nicht kommerzielle Zwecke verfolgen, dürfen Vervielfältigungen nach Absatz 1 folgenden Personen öffentlich zugänglich machen:
 1. einem bestimmt abgegrenzten Kreis von Personen für deren gemeinsame wissenschaftliche Forschung sowie.
 2. einzelnen Dritten zur Überprüfung der Qualität wissenschaftlicher Forschung.

Sobald die gemeinsame wissenschaftliche Forschung oder die Überprüfung der Qualität wissenschaftlicher Forschung abgeschlossen ist, ist die öffentliche Zugänglichmachung zu beenden.

- (6) Berechtigte nach den Absätzen 2 und 3 Nummer 1 dürfen Vervielfältigungen nach Absatz 1 mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung aufbewahren, solange sie für Zwecke der wissenschaftlichen Forschung oder zur Überprüfung wissenschaftlicher Erkenntnisse erforderlich sind.
- (7) Rechtsinhaber sind befugt, erforderliche Maßnahmen zu ergreifen, um zu verhindern, dass die Sicherheit und Integrität ihrer Netze und Datenbanken durch Vervielfältigungen nach Absatz 1 gefährdet werden.

Literaturverzeichnis

- Andresen, Melanie, Datengeleitete Sprachbeschreibung mit syntaktischen Annotationen, Eine Korpusanalyse am Beispiel der germanistischen Wissenschaftssprachen, Tübingen 2022.
- Andresen, Melanie et al., Vorzüge von Auszügen – Urheberrechtlich geschützte Texte in den digitalen Geisteswissenschaften (nach-)nutzen, Zeitschrift für digitale Geisteswissenschaften (ZfdG) im Erscheinen 2022.
- Bunk, Patrick, Text und Data Mining im Kontext von Smart Data – eine wirtschaftliche Perspektive in Smart-Data-Begleitforschung, Forschungszentrum Informatik (Hrsg.), Daten als Wirtschaftsgut, Berlin 2017, S. 30.
- de la Durantaye, Katharina; Raue, Benjamin, Urheberrecht und Zugang in einer digitalen Welt – Urheberrechtliche Fragestellungen des Zugangs für Gedächtnisinstitutionen und die Digital Humanities, RuZ – Recht und Zugang (RuZ) 2020, S. 83.
- Deutsche Forschungsgemeinschaft, Leitlinien zur Sicherung guter wissenschaftlicher Praxis Kodex, doi: <https://doi.org/10.1002/9783527679188.oth1>.
- Döhl, Frédéric, Digitaler Game Changer für die Gedächtnisinstitutionen? – Digital Humanities, die anstehende Schranke der Urheber- und Leistungsschutzrechte zugunsten der wissenschaftlichen Nachnutzung von Korpora bei Text und Data Mining und die unadressiert gebliebene Herausforderung des rechtmäßigen Zugangs, RuZ – Recht und Zugang (RuZ) 2020, S. 195.

- Dreier, Thomas; Schulze (Hrsg.), Urheberrechtsgesetz, Verwertungsgesellschaftengesetz, Kunsturhebergesetz, UrhDaG, 7. Aufl., München 2022 (zitiert: Bearbeiter in Dreier/Schulze, UrhG, 7. Auflage 2022).
- Dreier, Thomas; Schulze, Gernot (Hrsg.), Urheberrechtsgesetz, Verwertungsgesellschaftengesetz, Kunsturhebergesetz, UrhDaG, 6. Auflage 2018, München 2018 (zitiert: Bearbeiter in Dreier/Schulze, UrhG, 6. Auflage 2018).
- Ducato, Rossana; Strowe, Alain, Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to “Machine Legibility”, *International Review of Intellectual Property and Competition Law (IIC)* 2019, S. 649.
- Euler, Ellen, Open Access in der Wissenschaft und die Realitäten des Rechts, *RuZ – Recht und Zugang (RuZ)* 2020, S. 56.
- Gärtner, Markus; Kleinkopf, Felicitas; Andresen, Melanie; Hermann, Sibylle, Corpus Reusability and Copyright – Challenges and Opportunities, *Proceedings of the Workshop on Challenges in the Management of Large Corpora (CMLC-9)* 2021, Limerick 12. Juli 2021, S. 10; doi: <https://doi.org/10.14618/ids-pub-10470>.
- Geiger, Christophe; Frosio, Giancarlo; Bulanyenko, Oleksandr, Text and Data Mining in the Proposed Copyright Reform: Making the Eu Ready for an Age of Big Data? Legal Analysis and Policy Recommendations, *International Review of Intellectual Property and Competition Law (IIC)* 2018, S. 814.
- Grisse, Karina; Kaiser, Carola, Freiheit ab Unkenntlichkeit? – Die Bedeutung der (fehlenden) Wiedererkennbarkeit für das urheberrechtliche Vervielfältigungsrecht, *Zeitschrift für Urheber- und Medienrecht (ZUM)* 2021, S. 401.
- Grünberger, Michael, Bedarf es einer Harmonisierung der Verwertungsrechte und Schranken?, Ein Beitrag zur Entwicklung dogmatischer Bausteine eines umweltsensiblen Urheberrechts, *Zeitschrift für Urheber- und Medienrecht (ZUM)* 2015, S. 273.
- Hacker, Philipp, Ein Rechtsrahmen für KI-Trainingsdaten, *Zeitschrift für geistiges Eigentum (ZGE)* 2020, S. 239; doi: <https://doi.org/10.1628/zge-2020-0013>.
- Handig, Christian, Durch „freie kreative Entscheidungen“ zum europäischen urheberrechtlichen Werkbegriff, *Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil (GRUR Int.)* 2012, S. 973.
- Handke, Christian; Guibault, Lucie; Vallbb, Joan-Josep, Is Europe Falling Behind in Data Mining? Copyright’s Impact on Data Mining in Academic Research in Schmidt/Dobrevá (Hrsg.), *New avenues for electronic publishing in the age of infinite collections and citizen science*, Amsterdam 2015, doi: <https://doi.org/10.3233/978-1-61499-562-3-120>.
- Haucap, Justus; Loebert, Ina; Spindler, Gerald; Thorwarth, Susanne, Ökonomische Auswirkungen einer Bildungs- und Wissenschaftsschranke im Urheberrecht, Eine Studie im Auftrag des Bundesministeriums für Bildung und

- Forschung (BMBF), Düsseldorf 2016, abrufbar unter <https://www.econstor.eu/handle/10419/144535>.
- Hauk, Stefanie; Pflüger, Thomas, Die Umsetzung der DSM-RL in das Urheberrecht – Neujustierung der urheberrechtlichen Rahmenbedingungen für Hochschulen, Forschungs- und Kultureinrichtungen in der digitalen Welt, Zeitschrift für Urheber- und Medienrecht (ZUM) 2020, S. 383.
- Hoeren, Thomas; Sieber, Ulrich; Holznagel, Bernd; Altenhain, Karsten; Albrecht, Florian (Hrsg.), Handbuch Multimedia-Recht, Rechtsfragen des elektronischen Geschäftsverkehrs, Stand: Februar 2021 (55. EL), München 2021 (zitiert: Bearbeiter in Hoeren/Sieber/Holznagel, Multimediarecht Hdb).
- Jacke, Janina, Systematik unzuverlässigen Erzählens, Analytische Aufarbeitung und Explikation einer problematischen Kategorie, Berlin 2020; doi: <https://doi.org/10.1515/9783110659689>.
- Jannidis, Fotis; Kohle, Hubertus; Rehbein, Malte, Digital Humanities, Eine Einführung, Stuttgart 2017.
- Jotzo, Florian, Der Schutz großer Textbestände nach dem UrhG – Die Nutzbarmachung fremder Textbestände für die Forschung, RuZ – Recht und Zugang (RuZ) 2020, S. 128.
- Kleinkopf, Felicitas; Jacke, Janina; Gärtner, Markus, Text- und Data-Mining, Urheberrechtliche Grenzen der Nachnutzung wissenschaftlicher Korpora bei computergestützten Verfahren und digitalen Ressourcen, MMR Zeitschrift für IT-Recht und Recht der Digitalisierung (MMR) 2021, S. 196; doi: <https://doi.org/10.18419/opus-11445>.
- Kleinkopf, Felicitas; Pflüger, Thomas, Digitale Bildung, Wissenschaft und Kultur, Welcher urheberrechtliche Reformbedarf verbleibt nach Umsetzung der DSM-RL durch das Gesetz zum Urheberrecht im digitalen Binnenmarkt?, Zeitschrift für Urheber- und Medienrecht (ZUM) 2021, S. 643.
- Kuschel, Linda, Urheberrecht und Forschungsdaten, Ordnung der Wissenschaft (OdW) 2020, S. 43.
- Leventer, N. Orly, Google Book Search und vergleichendes Urheberrecht, Unter Heranziehung des deutschen und des US-amerikanischen Rechts, Baden-Baden 2012.
- Loewenheim, Ulrich (Hrsg.), Handbuch Urheberrecht, 3. Auflage, München 2020 (zitiert: Bearbeiter in Loewenheim UrhR-HdB).
- Maunz, Theodor; Dürig, Günter (Hrsg.), Grundgesetz, Kommentar, 91. Lieferung, München 2020 (zitiert: Bearbeiter in Maunz/Dürig, GG).
- Moltke, Bertram von, Das Urheberrecht an den Werken der Wissenschaft, Baden-Baden 1992.
- Nordemann, Axel; Nordemann, Jan Bernd (Hrsg.), Urheberrecht, Kommentar zum Urheberrechtsgesetz, zum Verlagsgesetz und zum Urheberrechtswahrnehmungsgesetz, Stuttgart 2018 (zitiert: Bearbeiter in Fromm/Nordemann, UrhG).

- Raue, Benjamin, Das Urheberrecht in der digitalen Wissen(schaft)sgesellschaft, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2017, S. 11.
- Ders., Text und Data Mining, Die neue Urheberrechtsschranke des § 60d UrhG, Computer und Recht (CR) 2017, S. 656; doi: <https://doi.org/10.9785/cr-2017-1008>.
- Ders., Rechtssicherheit für datengestützte Forschung, Zeitschrift für Urheber- und Medienrecht (ZUM) 2019, S. 684.
- Ders., Die Freistellung von Datenanalysen durch die neuen Text und Data Mining-Schranken (§§ 44b, 60d UrhG), Zeitschrift für Urheber- und Medienrecht (ZUM) 2021, S. 793.
- Raue, Benjamin; Schöch, Christof, Zugang zu großen Textkorpora des 20. und 21. Jahrhunderts mit Hilfe abgeleiteter Textformate – Versöhnung von Urheberrecht und textbasierter Forschung, RuZ – Recht und Zugang (RuZ) 2020, S. 118.
- Rieger, Sören, Der rechtliche Schutz wissenschaftlicher Datenbanken, Tübingen 2010.
- Schack, Haimo, Schutzgegenstand, „Ausnahmen oder Beschränkungen“ des Urheberrechts, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2021, S. 904.
- Schöch, Christof et al., Abgeleitete Textformate: Text und Data Mining mit urheberrechtlich geschützten Textbeständen, Zeitschrift für digitale Geisteswissenschaften (ZfdG) 2020; doi: https://doi.org/10.17175/2020_006.
- Schricker, Gerhard; Loewenheim, Ulrich (Hrsg.), Urheberrecht, Kommentar, 5. Auflage, München 2017 (zitiert: Bearbeiter in Schricker/Loewenheim, UrhG, 5. Auflage 2017).
- Schricker, Gerhard; Loewenheim, Ulrich (Hrsg.), Urheberrecht, UrhG, KUG, VGG, 6. Auflage, München 2020 (zitiert: Bearbeiter in Schricker/Loewenheim, UrhG, 6. Auflage 2020).
- Senftleben, Martin, Grundprobleme des urheberrechtlichen Dreistufentests, Gewerblicher Rechtsschutz und Urheberrecht Internationaler Teil (GRUR Int.) 2004, S. 200.
- Shen, Dan, Unreliability in Hühn/Meister/Pier/Schmid (Hrsg.), the living handbook of narratology, Hamburg, abrufbar unter <http://www.lhn.uni-hamburg.de/node/66.html>.
- Specht, Louisa, Die neue Schrankenregelung für Text und Data Mining und ihre Bedeutung für die Wissenschaft, Ordnung der Wissenschaft (OdW) 2018, S. 285.
- Specht-Riemenschneider, Louisa, Diktat der Technik, Regulierungskonzepte technischer Vertragsinhaltsgestaltung am Beispiel von Bürgerlichem Recht und Urheberrecht, Baden-Baden 2019.
- Spindler, Gerald, Europäisches Urheberrecht in der Informationsgesellschaft, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2002, S. 105.
- Ders., Text und Data Mining – urheber- und datenschutzrechtliche Fragen, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2016, S. 1112.

- Ders., Text- und Datamining im neuen Urheberrecht und in der europäischen Diskussion, *Zeitschrift für geistiges Eigentum (ZGE)* 2018, S. 273; doi: <https://doi.org/10.1628/zge-2018-0023>.
- Spindler, Gerald; Schuster, Fabian (Hrsg.), *Recht der elektronischen Medien*, 4. Auflage, München 2019 (zitiert: Bearbeiter in Spindler/Schuster, *Recht der elektronischen Medien*).
- Stieper, Malte, *Rechtfertigung, Rechtsnatur und Disponibilität der Schranken des Urheberrechts*, Tübingen 2009, abrufbar unter <http://gbv.eblib.com/patron/FullRecord.aspx?p=872801>.
- Wandtke, Artur-Axel, *Werkbegriff im Urheberrechts-Wissensgesellschafts-Gesetz*, *Neue Juristische Wochenschrift (NJW)* 2018, S. 1129.
- Ders., *Grundsätze der Richtlinie über das Urheberrecht im digitalen Binnenmarkt*, *Neue Juristische Wochenschrift (NJW)* 2019, S. 1841.
- Wandtke, Artur-Axel; Bullinger, Winfried (Hrsg.), *Praxiskommentar zum Urheberrecht*, 4. Auflage, München 2014 (zitiert: Bearbeiter in Wandtke/Bullinger, *UrhG*, 4. Aufl. 2014).
- Wandtke, Artur-Axel; Bullinger, Winfried (Hrsg.), *Praxiskommentar zum Urheberrecht*, 5. Auflage, München 2019 (zitiert: Bearbeiter in Wandtke/Bullinger, *UrhG*).
- Wilkinson, Mark D. et al., *The FAIR Guiding Principles for scientific data management and stewardship*, *Scientific data (Scientific data)* 2016, S. 160018.
- Wirth, Thomas, *Die Pflicht zur Löschung von Forschungsdaten – Urheber- und Datenschutzrecht im Widerspruch zu den Erfordernissen guter wissenschaftlicher Praxis?*, *Zeitschrift für Urheber- und Medienrecht (ZUM)* 2020, S. 585.

Das Rechtsgutachten ist im Kontext des vom Ministerium für Wissenschaft, Forschung und Kunst des Landes Baden-Württemberg geförderten Projektes "XSample" entstanden. Ziel des Projektes war es, die urheberrechtlichen Möglichkeiten bei der Forschung mit Text- und Data-Mining zu prüfen. Inhalt dieses Gutachtens sind die rechtliche Beurteilung der urheberrechtlichen Relevanz realer Anwendungsfälle aus dem Bereich der digitalen Geisteswissenschaften sowie die Analyse der Erlaubnis der einzelnen Arbeitsschritte.

Einen zentralen Gegenstand dieses Forschungsprozesses bildet die Sicherung der Anforderungen der guten wissenschaftlichen Praxis, womit die Frage einhergeht, ob die für das Text- und Data-Mining erstellten Korpora jedenfalls auszugsweise zu Forschungszwecken weitergegeben werden dürfen. Auf diese Weise soll eine gewisse Rechtssicherheit für Forschende erreicht werden, die sich mit Fragen nach der Reichweite der neuen Text- und Data-Mining-Erlaubnisse im Urheberrechtsgesetz - §§ 44b, 60d - konfrontiert sehen, um die Forschung auf urheberrechtlich geschützten Werken zukünftig in einem für die Forschung ausreichenden Umfang reproduzierbar zu gestalten.



FELICITAS KLEINKOPF

Wissenschaftliche Mitarbeiterin am Zentrum für Angewandte Rechtswissenschaft, Institut für Informations- und Wirtschaftsrecht am Karlsruher Institut für Technologie (KIT) sowie Rechtsreferendarin am Oberlandesgericht Karlsruhe.

In Kooperation mit



Universität Stuttgart



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST



ZAR

Zentrum für Angewandte Rechtswissenschaft



ISBN 978-1-914481-28-4



9 781914 481284