

IRDT PAPER SERIES Nr. 6

Die Text und Data Mining-Schranken und ihr Rahmen
für Textanalysen in den Digital HumanitiesKatharina Erler-Fridgen¹

Version 1.0 (02.06.2022), CC BY-SA 4.0.

Werden Informationen aus urheberrechtlich geschützten Texten² extrahiert und zu diesem Zweck Quellen, beispielsweise aus Datenbanken,³ gesammelt, aufbereitet und die Ergebnisse der Textanalyse sowie die Ausgangstexte schließlich aufbewahrt, so werden Vervielfältigungshandlungen oder Entnahmen vorgenommen, die einer urheberrechtlichen Gestattung bedürfen. Die Text und Data Mining-Schranken in § 44b UrhG und § 60d UrhG schaffen hierfür einen Rahmen, der im Folgenden erläutert werden soll.

¹ Die Verfasserin Dipl.-Jur. Katharina Erler-Fridgen ist wissenschaftliche Mitarbeiterin am Institut für Recht und Digitalisierung Trier bei Prof. Dr. Benjamin Raue (IRDT, Universität Trier) und arbeitet im interdisziplinären Forschungsprojekt Mining and Modeling Text (MiMoText, Universität Trier).

² Zu den Kriterien der urheberrechtlichen Schutzfähigkeit siehe *Erler-Fridgen*, Kriterien der urheberrechtlichen Schutzfähigkeit von Texten und Sammelwerken, IRDT PAPER SERIES Nr. 2.

³ Zum Schutz von Datenbanken und der Entnahme hieraus siehe *Erler-Fridgen*, Datenbanken als Quelle oder Ergebnis von Textanalysen – Datenbankwerkschutz und das Leistungsschutzrecht des Datenbankherstellers, IRDT PAPER SERIES Nr. 4.

Institut für Recht und Digitalisierung Trier (IRDT), Direktoren: Prof. Dr. Timo Hebler, Prof. Dr. Benjamin Raue, Prof. Dr. Peter Reiff, Prof. Dr. Antje von Ungern-Sternberg
Universität Trier, Campus II, Gebäude H, 54286 Trier, Behringstraße 21, 54296 Trier, <https://irdt.uni-trier.de>

Zitervorschlag:

Erler-Fridgen, Die Text und Data Mining-Schranken und ihr Rahmen für Textanalysen in den Digital Humanities, IRDT PAPER SERIES Nr. 6.

I. Hintergrund des Text und Data Mining in den Digital Humanities⁴

Text und Data Mining wird eingesetzt, um Informationen aus Texten zu gewinnen. **Verfahren** des Text und Data Mining können etwa das Information Retrieval⁵, die Named Entity Recognition⁶ oder das Topic Modeling⁷ sein.⁸ Beispielsweise können mit diesen Verfahren Themenaussagen in Sekundärliteratur identifiziert oder eine Extraktion von Figuren und Handlungsorten aus Romanen vorgenommen werden. Dabei können die zugrundeliegenden literarischen Quellen aus **Datenbanken** heraus eingesetzt worden sein. Außerdem sind zur Umsetzung der Informationsextraktion regelmäßig zahlreiche **Kopierschritte** notwendig: Zunächst muss häufig ein Scan vorgenommen werden oder eine bereits vorliegende digitale Ressource eingesetzt werden. Die dann vorliegende Datei (Bild/PDF-Datei) wird mittels Optical Character Recognition (OCR)⁹ oder dem Double Keying Verfahren¹⁰ in eine auslesbare Textdatei umgewandelt. In der üblicherweise hieran anschließenden Phase des Preprocessings werden die Texte oftmals korrigiert, vereinheitlicht, in ein geeignetes Datenformat transformiert und gegebenenfalls auch segmentiert oder gefiltert. Weiterhin kann die Datei in ein Annotationstool importiert, dort bearbeitet und dann exportiert werden. Schließlich können quantitative Verfahren wie die Named Entity Recognition oder Verfahren des maschinellen Lernens angewandt werden.

Für die Digital Humanities können im Bereich des Text und Data Mining unterschiedliche Arten von **Textquellen** interessant sein. Zum einen können für die Digital Humanities Textsammlungen hilfreich sein, die durch die Universitätsbibliothek etwa über Lizenzen bereitgestellt werden. Ein Beispiel hierfür wäre das Corpus of Contemporary American English (COCA)¹¹, das aus Sicht der Digital Humanities für die Analyse auch mit Verfahren des Text und Data Mining geeignet ist. Zum anderen gibt es im Internet zugängliche Textbestände wie beispielsweise das wissenschaftliche Blogportal Hypotheses.¹² Hier sind die Texte teilweise über Creative-Commons lizenziert. Im Fall der Sammlung literarischer Texte auf der Plattform Teatro Espanol del Siglo de Oro (TESO)¹³ sind

⁴ Dank gilt Prof. Dr. Christof Schöch und Dr. Maria Hinzmann (MiMoText, Universität Trier) für den Austausch zum Thema und die Bereitstellung und Diskussion von Beispielen.

⁵ Ignatow/Mihalcea, Text Mining, S. 137 ff.

⁶ Neumann, 5.3 Text-basiertes Informationsmanagement, in K.-U. Carstensen, Ch. Ebert, C. Ebert, S. Jekat, R. Klabunde, H. Langer (Hrsg.), Computerlinguistik und Sprachtechnologie, 3. Aufl., S. 596 ff.

⁷ Blei, Probabilistic Topic Models, <http://www.cs.columbia.edu/~blei/papers/Blei2012.pdf> (abgerufen am 20.07.2021).

⁸ Zu unterschiedlichen Verfahrensmethoden siehe Schöch, Quantitative Analyse, in Jannidis/Kohle/Rehbein (Hrsg.), Digital Humanities, S. 279 ff.

⁹ Verfahren zur Textdigitalisierung, in dem der Prozess der Texterfassung automatisiert vorgenommen wird, siehe <http://digitalhumanities.berkeley.edu/resources/digitization-workflows-scanning-ocr-and-audio-transcription> (abgerufen am 03.08.2021).

¹⁰ Double Keying Verfahren: Double Keying beschreibt ein Verfahren zur Textdigitalisierung, in dem der Prozess der Texterfassung manuell vorgenommen wird, siehe <http://www.dhmuseum.uni-trier.de/node/49> (abgerufen am 03.08.2021).

¹¹ <https://www.english-corpora.org/coca/> (abgerufen am 18.05.2022).

¹² <https://hypotheses.org/about-hypotheses> (abgerufen am 18.05.2022).

¹³ <https://www.nationallizenzen.de/angebote/nlproduct.2006-03-10.7433818373> (abgerufen am 18.05.2022).

weiterhin gemeinfreie Texte vorgehalten; der Download aller oder Teile des Bestands, etwa zur Erstellung einer Sammlung, werden in der Lizenz der Sammlung untersagt.

II. Rechtlicher Rahmen des Text und Data Mining

Grundsätzlich ist das Text und Data Mining als solches keine urheberrechtlich relevante Handlung.¹⁴ Semantische Informationen in Texten – mit Ausnahme von fiktionalen Geschichten – sind urheberrechtlich nicht geschützt.¹⁵ Die **Extraktion** von Informationen aus urheberrechtlich geschützten Quellen als solche ist daher nicht vom Zuweisungsgehalt des Urheberrechts umfasst.¹⁶ Auch wenn die Analyse mit technischen Mitteln erfolgt, ändert sich auch nichts daran, dass die Analyse und das Herausfiltern von Informationen aus einem Text nicht vom Zuweisungsgehalt des Urheberrechts erfasst ist.¹⁷

Selbst wenn die automatisierte Auswertung als solche keine urheberrechtlich relevante Handlung darstellt, so sind die Informationen regelmäßig in einer urheberrechtlich geschützten Hülle¹⁸ oder in einer geschützten Datenbank¹⁹ enthalten und es können urheberrechtlich relevante **Vervielfältigungen** (§ 16 UrhG) oder **Entnahmen** (§ 87a ff. UrhG) als Vorstufe, zur Durchführung oder im Nachgang von Textanalysen notwendig sein.²⁰ Denn insbesondere das Vervielfältigungsrecht in Art. 2 InfoSoc-RL (2001/29/EG) bzw. § 16 UrhG ist weit auszulegen und erfasst auch das vorübergehende Hochladen in den Arbeitsspeicher^{21, 22}

Um in diesen Fällen **Rechtssicherheit** zu gewähren, führte bereits die Text und Data Mining-Schranke für wissenschaftliche Forschung (§ 60d UrhG alte Fassung (a.F.), zur Umsetzung von Art. 3 DSM-RL ((EU) 2019/790)²³ siehe § 60d UrhG) für solche Handlungen im Umfeld des Mining eine Ausnahme ein.²⁴ Als Schrankenbestimmung erfüllte § 60d UrhG a.F. die Funktion, Handlungen des Text und Data Mining zustimmungsfrei möglich zu machen, sofern dies zum Zweck der wissenschaftlichen Forschung erfolgte²⁵ und die handelnde Person bereits rechtmäßigen Zugang zu

¹⁴ Erwägungsgrund 9 der 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 (DSM-RL).

¹⁵ *Raue*, ZUM 2019, 684, 685 mwN., siehe insbesondere BGH GRUR 1981, 352, 353 – Staatsexamensarbeit.

¹⁶ *Raue*, ZUM 2019, 684, 686.

¹⁷ *Raue*, GRUR 2017, 11, 13.

¹⁸ Zu den Kriterien der urheberrechtlichen Schutzfähigkeit siehe *Erler-Fridgen*, Kriterien der urheberrechtlichen Schutzfähigkeit von Texten und Sammelwerken, IRDT PAPER.SERIES Nr. 2.

¹⁹ Zum Datenbankschutz siehe *Erler-Fridgen*, Datenbanken als Quelle oder Ergebnis von Textanalysen – Datenbankschutz und das Leistungsschutzrecht des Datenbankherstellers, IRDT PAPER.SERIES Nr. 4.

²⁰ *Raue*, ZUM 2019, 684, 685, zu den nötigen Kopiervorgängen siehe oben I.

²¹ [BGH GRUR 2011, 418 Rn. 13 – Used Soft](#); *Schulze*, in Dreier/Schulze, Urheberrechtsgesetz, 7. Aufl. 2022, § 16 Rn. 13 mwN.

²² *Raue*, GRUR 2017, 11, 13.

²³ 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 (DSM-RL).

²⁴ *Raue*, GRUR 2017, 11, 13.

²⁵ Zum allgemeinen Text und Data Mining-Schranke siehe unten III.

den betroffenen Werken hatte.²⁶ Die bisher bestehende Schranke für das Text und Data Mining in der wissenschaftlichen Forschung in § 60d UrhG a.F. stützte sich auf Art. 5 Abs. 3 lit. a InfoSoc-RL (2001/29/EG)²⁷ und Art. 6 Abs. 2 lit. b Datenbank-RL²⁸ und war insofern eine Innovation des deutschen Gesetzgebers^{29,30} Mit ihr griff der deutsche Gesetzgeber der damals noch unsicheren Verabschiedung der DSM-RL ((EU) 2019/790) und der in deren Art. 3 DSM-RL-E³¹ vorgeschlagenen Text und Data Mining-Schranke für die wissenschaftliche Forschung vor.³²

Nunmehr enthält die jüngst umgesetzte³³ DSM-RL ((EU) 2019/790) in Art. 4 eine **allgemeine Text und Data Mining-Schranke** sowie in Art. 3 eine weitere TDM-Schranke zum Zwecke der wissenschaftlichen Forschung. Das UrhG beinhaltet nach der Umsetzung in § 44b und § 60d entsprechende Schrankenregelungen. Nun gibt es also zusätzlich eine allgemeine Text und Data Mining-Schranke (§ 44b UrhG in Umsetzung von Art. 4 DSM-RL ((EU) 2019/790)), die unter gesonderten Bedingungen das Text und Data Mining für alle freistellt, sofern rechtmäßiger Zugang zu den Texten besteht.³⁴ Die Bedeutung dieser beiden TDM-Schranken für die Qualität von Textanalysen ist wesentlich: Zwar können vorübergehende Vervielfältigungen bei Textanalysen³⁵ weiterhin auf § 44a UrhG gestützt werden.³⁶ Jedoch ist etwa für die Zusammenstellung, Normalisierung und Annotation eines Datenkorpus bei fundierten Textanalysen regelmäßig eine dauerhafte Speicherung notwendig, die einer über § 44a UrhG hinausgehenden urheberrechtlichen Gestattung bedarf.³⁷

III. Text und Data Mining § 44b UrhG

Die neu eingeführte **allgemeine Text und Data Mining-Schranke** in § 44b UrhG stellt das Text und Data Mining für alle frei und damit auch für kommerzielle Zwecke.³⁸ Diese allgemeine TDM-Schranke trägt dem Umstand Rechnung, dass Text und Data Mining-Verfahren nicht nur

²⁶ *Dreier*, in *Dreier/Schulze, Urheberrechtsgesetz*, 6. Aufl. 2018, § 60d Rn. 1; siehe auch *Dreier*, in *Dreier/Schulze, Urheberrechtsgesetz*, 7. Aufl. 2022, § 60d Rn. 2.

²⁷ 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 (Info-Soc-RL).

²⁸ Richtlinie 96/9/EG des Europäischen Parlaments und des Rates vom 11. März 1996 über den rechtlichen Schutz von Datenbanken (Datenbank-RL).

²⁹ *Spindler*, CR 2019, 277 Rn. 2.

³⁰ *Dreier*, in *Dreier/Schulze, Urheberrechtsgesetz*, 7. Aufl. 2022, § 60d Rn. 2.

³¹ COM(2016) 593 final.

³² *Raue*, ZUM 2021, 793, 794.

³³ Gesetz zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, in der Fassung der Bekanntmachung vom 04. Juni 2021, BGBl. I S. 1204 ff.

³⁴ *Heesen/Jüngels*, RuZ 2021, 45, 46 f.; zur Voraussetzung des rechtmäßigen Zugangs siehe unten IV.

³⁵ Etwa solche, die lediglich vorübergehend in den Arbeitsspeicher eines Computers hochgeladen werden: *Raue*, ZUM 2021, 793, 795.

³⁶ *Raue*, ZUM 2021, 793, 795.

³⁷ *Raue*, ZUM 2021, 793, 795; siehe insbesondere *Raue*, IIC 2018, 379, 381: vorbereitende Handlungen wie die Digitalisierung analoger Quellen oder die Normalisierung und Annotierung des Korpus verursachen üblicherweise eine längere Speicherdauer der Vervielfältigungen als sie von Art. 5 Abs. 1 InfoSoc-RL/§ 44a UrhG freigestellt ist.

³⁸ *Heesen/Jüngels*, RuZ 2021, 45, 46 f.

bedeutende Instrumente für die wissenschaftliche Forschung sind, sondern auch bei privaten sowie staatlichen Institutionen Einsatz finden.³⁹ Sie erfasst Vervielfältigungen von Werken und Datenbankwerken und ist außerdem auf Leistungsschutzrechte wie etwa das Leistungsschutzrecht des Datenbankherstellers⁴⁰ (siehe § 87c Abs. 1 Nr. 4 UrhG) sowie durch den Verweis auf Schranken für Werke beispielsweise auch auf das Leistungsschutzrecht für wissenschaftliche Ausgaben (§ 70 UrhG)⁴¹ anwendbar.⁴²

Zu diesem Zweck definiert sie **Text und Data Mining** in Absatz 1 als „die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen“. Diese **Definition** entspricht im Wesentlichen der Definition des Text und Data Minings in Art. 2 DSM-RL ((EU) 2019/790).⁴³ Sie stellt jedoch klar, dass auch die Analyse von einzelnen Werken unter das Text und Data Mining fallen kann.⁴⁴ Generell ist es hiernach möglich, ein Korpus zum Zwecke des Text und Data Mining zu erstellen, wie es § 60d UrhG a.F. ausdrücklich regelte.⁴⁵ Digitale Werke in der Definition sind dabei nach der Begründung des Gesetzesentwurfs solche, die von vornherein in digitaler Form vorliegen, wohingegen digitalisierte Werke erst im Zuge des Text und Data Mining digitalisiert werden.⁴⁶ Das heißt, auch die Digitalisierung von analogen Werken wird als Vorstufe der Text und Data Analyse in § 44d Abs. 2 S. 1 UrhG erlaubt.⁴⁷ Nicht erfasst ist hingegen die Digitalisierung von Texten, die allein darauf abzielt, ein digitales Archiv zu erstellen.⁴⁸

Zulässig sind nach § 44b Abs. 2 UrhG Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Dieses Kriterium des **rechtmäßigen Zugangs** rechtfertigt die Schranke und ermöglicht es den Rechteinhabern, den Zugang zu ihren Werken zu kontrollieren und vergütungspflichtig zu stellen.⁴⁹ Rechtmäßiger Zugang liegt etwa bei entsprechenden vertraglichen Vereinbarungen zwischen Rechteinhabern und Nutzenden vor, besteht aber auch bei frei im

³⁹ *Raue*, ZUM 2021, 793, 795 unter Verweis auf Erwägungsgrund 18 S. 1 DSM-RL.

⁴⁰ Hierzu ausführlich *Erler-Fridgen*, Datenbanken als Quelle oder Ergebnis von Textanalysen – Datenbankwerkschutz und das Leistungsschutzrecht des Datenbankherstellers, IRDT PAPER.SERIES Nr. 4.

⁴¹ Zum Leistungsschutzrecht für wissenschaftliche Ausgaben siehe: *Erler-Fridgen*, Die Nutzung wissenschaftlicher Ausgaben für Textanalysen, IRDT PAPER.SERIES Nr. 1.

⁴² *Raue*, ZUM 2021, 793, 796; für § 70 UrhG siehe *Loewenheim*, in Schricker/Loewenheim, Urheberrecht, 6. Auflage 2020, § 70 Rn. 9.

⁴³ Art. 2 Nr. 2 DSM-RL: Text und Data Mining bezeichnet eine Technik, für die automatisierte Analyse von Texten und Daten in digitaler Form, mit deren Hilfe Informationen unter anderem – aber nicht ausschließlich – über Muster, Trends und Korrelationen gewonnen werden können.

⁴⁴ *Raue*, ZUM 2020, 172, 172.

⁴⁵ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁴⁶ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁴⁷ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁴⁸ *Raue*, ZUM 2021, 793, 795 unter Hinweis auf Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁴⁹ *Raue*, ZUM 2019, 684, 686.

Internet verfügbaren Gegenständen oder solchen, die im Wege des Open Access frei zugänglich sind.⁵⁰

Umfasst werden Vervielfältigungen für das Text und Data Mining.⁵¹ Dabei ist zu beachten, dass das im Diskussionsentwurf des BMJV enthaltene Erforderlichkeitskriterium (es lautete in § 44b Abs. 2: „... sofern sie für das Text und Data Mining erforderlich sind ...“) entfallen ist⁵² und damit keine enge Auslegung der Verbindung zwischen Vervielfältigung und Text und Data Mining vorgenommen werden muss.⁵³ Dies ermöglicht es nun, den **Zweck der Vervielfältigung** für das Text und Data Mining wie es auch in Art. 4 Abs. 1 DSM-RL ((EU) 2019/790) („zum Zweck des Text und Data Mining“) – zu Gunsten der angestrebten Rechtssicherheit für die Forschung durch die Normen – breiter als im ursprünglichen Entwurf auszulegen.⁵⁴

Was die **Nachnutzbarkeit** und **Aufbewahrung** der erzeugten Vervielfältigungsstücke angeht, so sieht § 44b UrhG vor, dass diese zu löschen sind, wenn sie für das Text und Data Mining nicht mehr erforderlich sind. Die umzusetzende DSM-RL ((EU) 2019/790) spricht in Art. 4 Abs. 2 in diesem Kontext davon, dass Vervielfältigungen und Entnahmen so lange aufbewahrt werden dürfen, wie sie für Zwecke des TDM notwendig sind. Eine solche Löschpflicht nach Abschluss der Forschungsarbeiten war in der Vorgängervorschrift zum Text und Data Mining in der Forschung in § 60d UrhG a.F. ebenfalls enthalten, ist jedoch im neuen § 60d UrhG nicht mehr vorgesehen. Damit ist in § 44b UrhG keine dauerhafte Möglichkeit zur Aufbewahrung geregelt,⁵⁵ sondern eine Löschpflicht nach Erfüllung ihres Zweckes, der Erforderlichkeit für das Text und Data Mining, vorgesehen.

Es soll den Rechteinhabern bei der kommerziellen Forschung in § 44b UrhG möglich sein, Nutzungen vorzubehalten nach § 44b Abs. 3 S. 1 UrhG (in Umsetzung von Art. 4 Abs. 3 DSM-RL ((EU) 2019/790)). Das bedeutet, dass Mining-Handlungen von Rechteinhabern aktiv verboten werden müssen.⁵⁶ Diese Regelung verschafft Rechteinhabern die Möglichkeit, weiterhin Lizenzen für das Text und Data Mining einzuräumen.⁵⁷ Nach § 44b Abs. 3 S. 2 UrhG ist ein solcher **Nutzungsvorbehalt** bei online zugänglichen Werken jedoch nur dann wirksam, wenn er in maschinenlesbarer Form erfolgt. Dies setzt auf der Voraussetzung in Art. 4 Abs. 3 DSM-RL ((EU) 2019/790) auf, dass die Rechteinhaber ausdrücklich und in angemessener Weise den Nutzungsvorbehalt vorsehen müssen.⁵⁸ Der Nutzungsvorbehalt nach § 44b Abs. 3 S. 1 UrhG wirkt jedoch lediglich für die Zukunft (ex nunc) und berührt Textanalysen nicht, die sich auf die

⁵⁰ Raue, ZUM 2021, 793, 796 sowie Schack, GRUR 2021, 904, 907.

⁵¹ Siehe auch Raue, ZUM 2019, 684, 686.

⁵² Siehe § 44b Abs. 2 Disk-E des Bundesministeriums der Justiz und für Verbraucherschutz vom 15.01.2020.

⁵³ Raue, ZUM 2020, 172, 173.

⁵⁴ Zum Disk-E des BMJV vom 15.01.2020 siehe noch Raue, ZUM 2020, 172, 173.

⁵⁵ Heesen/Jüngels, RuZ 2021, 45, 46.

⁵⁶ Heesen/Jüngels, RuZ 2021, 45, 46.

⁵⁷ Raue, ZUM 2021, 793, 796; Erwägungsgrund 18 S. 2 DSM-RL.

⁵⁸ Hierzu auch Raue, ZUM 2020, 172, 173.

Schranken nach § 44a UrhG (bei vorübergehenden Vervielfältigungen) oder auf § 60d UrhG stützen.⁵⁹

Für das Text und Data Mining sieht § 44b UrhG **keine Vergütung** vor, da dies nach der Begründung des Gesetzesentwurfes der Rechtssicherheit diene.⁶⁰ Denn diese Rechtssicherheit könne nicht erreicht werden, wenn ansonsten in Zukunft immer ermittelt werden müsse, ob eine flüchtige Vervielfältigung (dann in jedem Fall vergütungsfrei nach § 44a UrhG) oder eine sonstige Vervielfältigung vorliege (anderenfalls zweifelhaft, ob vergütungsfrei).⁶¹ Diese Abgrenzungsproblematik entfalle demnach durch die Vergütungsfreistellung.⁶²

IV. Text und Data Mining zum Zwecke der wissenschaftlichen Forschung

§ 60 d UrhG setzt Art. 3 DSM-RL ((EU) 2019/790) um und modifiziert die Regelungen in § 44b UrhG (Art. 4 DSM-RL) für das allgemeine Text und Data Mining, wenn es für die **wissenschaftliche Forschung** eingesetzt wird.⁶³ Das bedeutet, dass nach § 60d UrhG Forschende, Forschungsorganisationen und Einrichtungen des Kulturerbes urheberrechtlich geschützte Gegenstände automatisiert analysieren dürfen, wenn sie rechtmäßigen Zugang zu diesen haben.⁶⁴ Bereits seit 2018 existierte in § 60d UrhG a.F. eine auf Grundlage der InfoSoc-RL⁶⁵ eingeführte deutsche Text und Data Mining-Schranke für nicht kommerzielle – wissenschaftliche – Zwecke.⁶⁶ Weitergehende Regelungen⁶⁷ aus § 60d UrhG a.F. werden auch nach Umsetzung von Art. 3 DSM-RL ((EU) 2019/790) beibehalten, was unionsrechtlich ausdrücklich gestattet ist.⁶⁸

⁵⁹ *Raue*, ZUM 2021, 793, 796.

⁶⁰ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88. Art. 4 DSM-RL enthält dazu keine Aussage, es fehlt auch an einer ausdrücklichen Option für die Mitgliedsstaaten eine Vergütungspflicht vorzusehen; ausführlich dazu auch *Raue*, ZUM 2020, 172, 173.

⁶¹ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁶² Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 88.

⁶³ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 95.

⁶⁴ *Raue*, ZUM 2021, 793, 797.

⁶⁵ 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.

⁶⁶ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 95.

⁶⁷ Insbesondere wird in § 60d UrhG-E der Berechtigtenkreis über Forschungseinrichtungen oder Einrichtungen des Kulturerbes hinaus u.a. auf Einzelforscher erstreckt und die öffentliche Zugänglichmachung an Forschergruppen aufgeführt, *Raue*, ZUM 2020, 172, 173.

⁶⁸ Art. 5 Abs. 3 lit. a InfoSoc-RL in der Neufassung durch Art. 24 Abs. 2 lit. b DSM-RL in Verbindung mit Art. 25 DSM-RL gestattet dies nach der Gesetzesentwurfsbegründung, siehe Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 95.

Für die **Definition** des Text und Data Mining verweist § 60d Abs. 1 UrhG auf § 44b UrhG und legt zudem fest, dass in Abgrenzung zu den allgemeinen Regelungen in § 44b UrhG für TDM zum Zwecke der wissenschaftlichen Forschung die Bestimmungen des § 60d Abs. 2 - 6 UrhG maßgeblich sind. § 44b Absatz 1 UrhG definiert **Text und Data Mining** wie beschrieben als „die automatisierte Analyse von einzelnen oder mehreren digitalen oder digitalisierten Werken, um daraus Informationen insbesondere über Muster, Trends und Korrelationen zu gewinnen.“ Zulässig sind Vervielfältigungen **zum Zweck des** Text und Data Minings, sodass neben der automatisierten Analyse, die Quellen auch zuvor gespeichert, normalisiert, annotiert oder auf sonstige Weise bearbeitet (§ 23 Abs. 3 UrhG) werden können.⁶⁹

Voraussetzung für die Zulässigkeit von Vervielfältigungen für TDM ist wie in § 44b Abs. 2 S. 1 UrhG der **rechtmäßige Zugang** zu den zugrundeliegenden Werken. Dabei gilt der Zugang zu frei im Internet verfügbaren Inhalten als rechtmäßig.⁷⁰ Außerdem kann hierunter auch nach Erwägungsgrund 14 der DSM-RL ((EU) 2019/790) der Zugang durch eine Open-Access Strategie, durch vertragliche Vereinbarungen zwischen Rechteinhabern und Forschungsinstitutionen wie Abonnements oder sonstige rechtmäßige Mittel fallen. Außerdem besteht rechtmäßiger Zugang zu Gegenständen, die nach Erwägungsgrund 14 S. 4 der DSM-RL ((EU) 2019/790) frei im Internet verfügbar sind sowie auch nach dem Erwerb oder der Fernleihe von analogen Medien.⁷¹ Beispielsweise werden in Projekten digitale Zeitschriften oder E-Books, die über die Universitätsbibliotheken entsprechend lizenziert sind, eingesetzt. Auch sonstige, rechtmäßig zugängliche Datenbanken, können die Quelle für digitale Ressourcen sein.

Zur Vervielfältigung **berechtigt** sind nach Abs. 2 **Forschungsorganisationen**, die hier legaldefiniert werden als „Hochschulen, Forschungsinstitute oder sonstige Einrichtungen, die wissenschaftliche Forschung betreiben, sofern sie 1) nicht kommerzielle Zwecke verfolgen, 2) sämtliche Gewinne in die Forschung reinvestieren oder 3) im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind“. Der Begründung des Gesetzesentwurfs nach entspricht diese Definition inhaltlich derjenigen in Art. 2 Nr. 1 DSM-RL ((EU) 2019/790) und ist lediglich in ihrer Formulierung gestrafft.⁷² Dem Wortlaut nach („oder“) genügt es, wenn eine von den drei genannten Voraussetzungen erfüllt ist.⁷³

Der Betrieb von „**wissenschaftlicher Forschung**“ in dieser Definition bezieht sich nach Erwägungsgrund 12 S. 2 DSM-RL sowohl auf die Naturwissenschaften als auch auf die

⁶⁹ Raue, ZUM 2021, 793, 797 und 798.

⁷⁰ Erwägungsgrund 18 UAbs. 2 S. 1 DSM-RL, Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96; wobei die Art der Nutzung (Download oder anderes) bei im Internet frei verfügbaren Inhalten im Rahmen der Frage nach dem rechtmäßigen Zugang wohl zunächst zu keiner Änderung führen dürfte, die übrigen Voraussetzungen des TDM müssen berücksichtigt werden; insbesondere muss wissenschaftliche Forschung bezweckt werden.

⁷¹ Raue, ZUM 2021, 793, 800; Spindler CR 2019, 277, 280.

⁷² Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96.

⁷³ So auch: Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96.

Geisteswissenschaften.⁷⁴ Die in der DSM-RL ((EU) 2019/790) ausdrücklich genannte Lehrtätigkeit (auch in Verbindung mit wissenschaftlicher Forschung) wird in der Definition von Forschungsorganisationen nicht aufgeführt.⁷⁵ Jedoch betreiben mit **Lehre** befasste Institutionen ganz regelmäßig auch Forschung und etwa Fachhochschulen werden zutreffenderweise wohl auch unter die Definition gefasst.⁷⁶ Ist die Forschung jedoch gänzlich der Lehre untergeordnet, etwa bei Schulen, befindet sich dies außerhalb der Definition der Richtlinie.⁷⁷

Für die **Qualifikation als Forschungsorganisation** ist es jedoch unerheblich, welche Rechtsform oder Struktur die Forschungseinrichtung prägt.⁷⁸ Jedenfalls darf kein kommerzieller Zweck verfolgt werden (Nr. 1), der eine Gewinnerorientierung der Einrichtung darstellen würde.⁷⁹ Zur Annahme einer Forschungsorganisation kommt man hingegen auch, wenn die Einrichtung ihre Gewinne vollständig in die Forschung reinvestiert (Nr. 2).⁸⁰ Auch werden Forschungseinrichtungen von der Schranke umfasst, die im Rahmen eines staatlich anerkannten Auftrags im öffentlichen Interesse tätig sind (Nr. 3). Typisch für einen staatlich anerkannten **Auftrag im öffentlichen Interesse** ist beispielsweise die Finanzierung durch die öffentliche Hand oder Bestimmungen zum öffentlichen Interesse in Rechtsvorschriften oder Aufträgen der öffentlichen Hand.⁸¹

Nicht berechtigt sind hingegen nach § 60d Abs. 2 S. 2 UrhG 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. Bei sogenannten Public Private Partnerships und damit verbundenen Kooperation mit privaten Personen beim Text und Data Mining soll es demnach zwar möglich sein, sich auf die Text und Data Mining Schranke zu berufen, dies aber entsprechend eingeschränkt werden, wenn ein bestimmender Einfluss oder ein bevorzugter Zugang vorliegt.⁸² Beispielsweise kann ein bestimmender Einfluss eines privaten Unternehmens gegeben sein, wenn es aufgrund struktureller Gegebenheiten als Anteilseigner Kontrolle ausüben kann und hierdurch auch bevorzugten Zugang zu Forschungsergebnissen erhält.⁸³

Ferner berechtigt sind nach Absatz 3 **Bibliotheken und Museen**, wenn sie öffentlich zugänglich sind, sowie Archive und Einrichtungen im Bereich des Film- oder Tonerbes (Kulturerbe-

⁷⁴ Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96.

⁷⁵ Art. 2 Nr. 1 DSM-RL; *Raue*, ZUM 2020, 172, 173.

⁷⁶ *Raue*, ZUM 2020, 172, 173; *Spindler*, CR 2019, 277 Rn. 6.

⁷⁷ *Spindler*, CR 2019, 277 Rn. 6.

⁷⁸ *Raue*, ZUM 2021, 793, 800 unter Verweis auf Erwägungsgrund 12 S. 4 DSM-RL.

⁷⁹ *Raue*, ZUM 2021, 793, 800: siehe Art. 2 Nr. 1 lit. a DSM-RL.

⁸⁰ Zur dabei zulässigen Preisbildung für Forschungsergebnisse siehe *Raue*, ZUM 2021, 793, 800 mwN.

⁸¹ Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96; Erwägungsgrund 12 S. 6 DSM-RL.

⁸² Erwägungsgrund 11 DSM-RL, Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96; siehe auch *Spindler*, CR 2019, 277 Rn. 10.

⁸³ Erwägungsgrund 12 S. 7 DSM-RL, Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 96.

Einrichtungen). Diese Legaldefinition setzt die neue Definition der Kulturerbe-Einrichtungen aus Art. 2 Nr. 3 der DSM-RL ((EU) 2019/790) um, der diese umreißt als „eine öffentlich zugängliche Bibliothek oder Museum, Archiv oder eine im Bereich des Film- oder Tonerbes tätige Einrichtung“. ⁸⁴ In Abgrenzung zu den nach § 60d Abs. 2 S. 2 UrhG berechtigten Forschungsorganisationen, handelt es sich bei den Kulturerbe-Einrichtungen nicht um Forschungsinstitute oder andere Forschungsorganisationen, sondern beispielsweise um Universitätsbibliotheken und sonstige Bibliotheken, jedoch unter der Voraussetzung des öffentlichen Zugangs zu diesen Bibliotheken. ⁸⁵ Rein privat zugängliche Bibliotheken sind damit beispielsweise im Umkehrschluss nicht erfasst. ⁸⁶

Weitergehend als Art. 3 DSM-RL ((EU) 2019/790) berechtigt Absatz 3 (wie schon die Vorgängervorschrift § 60d UrhG a.F. s.o.) auch Einzelforscher, sofern sie nicht kommerzielle Zwecke verfolgen. Die Berechtigung von Einzelforschern geht über die Umsetzung von Art. 3 DSM-RL hinaus und kann aber über Art. 25 DSM-RL und Art. 5 III lit. a InfoSoc-RL gefasst werden. ⁸⁷

Die Berechtigten dürfen die Vervielfältigungen auch einem **bestimmten Personenkreis** für bestimmte Zwecke öffentlich zugänglich machen: Nach § 60d Abs. 4 Nr. 1 UrhG soll eine gemeinsame Forschung in Forschendengruppen ermöglicht werden und zu diesem Zweck die öffentliche Zugänglichmachung von Vervielfältigungen an einen bestimmt abgegrenzten Personenkreis gestattet sein. ⁸⁸ Zur Überprüfung der Qualität von wissenschaftlicher Forschung erlaubt § 60d Abs. 4 Nr. 2 UrhG es, einzelnen Dritten die Vervielfältigungen öffentlich zugänglich zu machen. Nach Abschluss der Forschungsarbeiten bzw. der Überprüfung der Qualität wissenschaftlicher Forschung, ist die öffentliche Zugänglichmachung nach § 60d Abs. 4 UrhG zu beenden.

Die öffentliche Zugänglichmachung zur **Qualitätsüberprüfung** war bereits in der Vorgängervorschrift § 60d UrhG a.F. enthalten und deren Fortgeltung beruht wie andere weitergehende Regelungen (s.o.) auf Art. 25 und Erwägungsgrund 5 S. 1 DSM-RL ((EU) 2019/790). ⁸⁹ Damit war bereits in der Vorgängerregelung während des Peer-Review-Verfahren vor einer Veröffentlichung die öffentliche Zugänglichmachung zur Qualitätssicherung an einzelne

⁸⁴ *Dreier*, in *Dreier/Schulze, Urheberrechtsgesetz*, 7. Aufl. 2022, § 60d Rn. 6; *de la Durantaye/Kuschel*, ZUM 2021, 785, 786.

⁸⁵ Näher hierzu *Raue*, ZUM 2021, 793, 801.

⁸⁶ Anderes gelte nach dem deutschen Gesetzeswortlaut für Archive und Einrichtungen des Film- oder Tonerbes: *Raue*, ZUM 2021, 793, 801.

⁸⁷ *Raue*, ZUM 2020, 172, 174; so auch *Heesen/Jüngels*, RuZ 2021, 45, 48.

⁸⁸ Die Regelung des § 60d Abs. 4 UrhG ist an dieser Stelle nicht ganz klar, denn „öffentliche Zugänglichmachung“ im Sinne des § 19a UrhG/Art. 3 Abs. 2 Info-Soc-RL (2001/29/EG) wird nach der Rechtsprechung des EuGH an einen unbegrenzten Personenkreis gerichtet, *Spindler*, CR 2019, 277, 280 mit Hinweis u.a. auf [EuGH C-527/15, ECLI:EU:C:2017:300, GRUR 2017, 610 Rn. 32 – Stichting Brein](#); Filmspeler. In § 60d Abs. 4 UrhG wird hingegen explizit auf einen begrenzten Personenkreis abgestellt; hierzu: *Raue*, CR 2017, 656, 660 sowie *Raue*, ZUM 2021, 793, 799; gegebenenfalls könne auch ein Zugänglichmachen vorliegen, auch wenn es sich um ein Zugänglichmachen im Sinne von § 19a UrhG handelt: *Spindler*, CR 2019, 277, 280.

⁸⁹ Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 97.

Dritte möglich.⁹⁰ Eine – ggf. über diesen rechtlichen Rahmen hinausgehende – Möglichkeit zur Qualitätsüberprüfung ist (nicht nur) aus Sicht der Digital Humanities ein wesentliches Desiderat: Beispielsweise wäre die Bereitstellung von Textteilen zur Qualitätsüberprüfung von Verfahren wie etwa der Named Entity Recognition aus Sicht der Digital Humanities ein hilfreicher Beleg für die Verlässlichkeit der Ergebnisse der Analyseverfahren.⁹¹

Neu ist, dass Berechtigte – nicht hingegen Einzelforscher ohne Anbindung an eine Forschungsinstitution – nach § 60d Abs. 5 UrhG die Vervielfältigungen mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung **aufbewahren** dürfen, solange sie für die **Zwecke der wissenschaftlichen Forschung** oder zur **Überprüfung der wissenschaftlichen Qualität erforderlich** sind. In der Vorgängervorschrift § 60d Abs. 3 UrhG a.F. war zuvor eine Löschpflicht nach Abschluss der Forschungsarbeiten vorgesehen (s.o.).⁹² Nach dieser Regelung ist es nun, anders als zuvor, möglich, dass die Forschungsorganisationen die Korpora selbst aufbewahren können, solange dies erforderlich ist. Das Merkmal „erforderlich“ ist an dieser Stelle enger als die Anforderungen der DSM-Richtlinie, nach der eine Bindung der Speicherung an den Zweck der wissenschaftlichen Forschung sowie auch an die Überprüfung der wissenschaftlichen Erkenntnisse ausreicht.⁹³ Demnach ist auch hier die Dauer der Aufbewahrung der Forschungsergebnisse eingeschränkt – nur bis zu dieser Erforderlichkeitsgrenze – geregelt. Jedenfalls zehn Jahre lang kann die Aufbewahrung zur **Überprüfung** wissenschaftlicher Erkenntnisse nach den Grundsätzen guter wissenschaftlicher Praxis für erforderlich gehalten werden.⁹⁴ Die Erforderlichkeit einer längeren Aufbewahrungsdauer für die **wissenschaftliche Forschung** – gestützt auf die Wissenschaftsfreiheit – müsste seitens der Forschenden etwa unter Verweis auf Anschlussforschung ausreichend plausibel gemacht werden.⁹⁵ Nach der Begründung des Gesetzesentwurfs kann auch eine dauerhafte Speicherung erforderlich und damit zulässig sein, insbesondere, wenn sie durch die Kulturerbe-Einrichtungen und nicht die Forschungseinrichtung selbst erfolgt.⁹⁶ Demnach und ausweislich seines Wortlauts scheint § 60 d Abs. 5 UrhG die Option der Vorgängervorschrift § 60d Abs. 3 S. 2 UrhG a.F. zu eröffnen, die Archivierung auch bei einer Kulturerbe-Einrichtung zentral vorzunehmen.⁹⁷ Voraussetzung der Aufbewahrung sind angemessene Sicherheitsvorkehrungen

⁹⁰ Zum alten Recht: *Raue*, CR 2017, 656, 695, Gesetzesentwurfsbegründung BT-Drucks. 18/12329, 41.

⁹¹ Hierzu ausführlich *Erler-Fridgen*, Die Präsentation von Textteilen als Ergänzung von Textanalysen, IRDT PAPER.SERIES Nr. 3.

⁹² *Raue*, ZUM 2020, 172, 174.

⁹³ *Raue*, ZUM 2020, 172, 174.

⁹⁴ *Raue*, ZUM 2021, 793, 799 nach Erörterung etwa von: DFG, Leitlinien zur Sicherung guter wissenschaftlicher Praxis, Kodex, 2019, Leitlinie 12, 13, insbesondere 17; Argumentation für eine längere Aufbewahrungszeit: *Heesen/Jüngels*, RuZ 2021, 45, 50.

⁹⁵ *Raue*, ZUM 2021, 793, 799: Forschende haben hier eine Einschätzungsprärogative, die gerichtlich lediglich eingeschränkt auf Missbrauch überprüft werden kann; bereits zur Definition des Forschungsprojektes durch die Forschenden unter § 60d UrhG a.F. siehe *Stieper*, in Schrickler/Loewenheim, Urheberrecht, 6. Auflage 2020, § 60d Rn. 17 und *Raue*, CR 2017, 656, 659.

⁹⁶ Gesetzesentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 97.

⁹⁷ Zu den Vorteilen der zentralen Archivierung siehe *Kleinkopf/Jacke/Gärtner*, MMR 2021, 196, 197, 198.

gegen unbefugte Benutzung nach § 60d Abs. 5 UrhG. Es wird aus diesem Grund in der Literatur empfohlen, diese Daten bei Forschungsdatenrepositorien aufzubewahren, die institutionell dafür Sorge tragen können, dass diese Sicherheitsstandards eingehalten werden.⁹⁸

Die Nutzung zum Text und Data Mining für die wissenschaftliche Forschung wird durch die Änderung des § 60h Abs. 2 UrhG **vergütungsfrei** gestellt. Denn in Erwägungsgrund 17 S. 2 DSM-RL ((EU) 2019/790) ist festgehalten, dass das Text und Data Mining zu Zwecken der wissenschaftlichen Forschung künftig entsprechend ausgestaltet werden soll.⁹⁹ Dies entspricht auch der zugrundeliegenden Wertung der Voraussetzung des rechtmäßigen Zugangs, der eine ausreichende Möglichkeit der Rechteinhaber zur vergütungspflichtigen Ausgestaltung des Zugangs bietet.¹⁰⁰ Schließlich können **vertragliche Einschränkungen** nach § 60g Abs. 1 UrhG und technische Schutzmaßnahmen nach § 95b Abs. 1 Nr. 11 und Abs. 3 Nr. 5 UrhG die Schranke des § 60d UrhG nicht beeinträchtigen.¹⁰¹ Der Vorrang der Schranke § 60d UrhG¹⁰² nach § 60g gilt nach § 137o UrhG jedoch nicht für Verträge, die vor dem 01. März 2018 geschlossen wurden. Neben der Wissenschaftsschranke § 60d UrhG können sich Forschende auch auf die Schranke für vorübergehende Vervielfältigungshandlungen nach § 44a UrhG sowie auf die allgemeine Text und Data Mining-Schranke nach § 44b UrhG berufen.¹⁰³

V. Ergebnis

Eine neu eingeführte allgemeine Text und Data Mining-Schranke in § 44b UrhG stellt das Text und Data Mining für alle und auch für kommerzielle Zwecke frei. Zulässig sind nach § 44b Abs. 2 S. 1 UrhG Vervielfältigungen von rechtmäßig zugänglichen Werken für das Text und Data Mining. Erstellte Vervielfältigungen sind jedoch nach § 44b Abs. 2 UrhG zu löschen, wenn sie für das Text und Data Mining nicht mehr erforderlich sind. Das Text und Data Mining ist vergütungsfrei, es können jedoch Nutzungen durch die Rechteinhaber nach § 44b Abs. 3 UrhG vorbehalten werden.

§ 60 d UrhG setzt Art. 3 DSM-RL ((EU) 2019/790) um und stellt das Text und Data Mining frei, wenn es für die wissenschaftliche Forschung eingesetzt wird. Voraussetzung für die Zulässigkeit von Vervielfältigungen für TDM ist wie in § 44b Abs. 2 S. 1 UrhG der rechtmäßige Zugang zu den zugrundeliegenden Werken. Berechtigt zur Durchführung von TDM sind hiernach Forschungsorganisationen und ihre Mitglieder (Abs. 2) sowie Bibliotheken, Museen und

⁹⁸ *Raue*, ZUM 2021, 793, 799; Erwägungsgrund 15 S. 3 DSM-RL führt diesbezüglich sogenannte vertrauenswürdige Stellen ein, die der deutsche Gesetzgeber entsprechend Erwägungsgrund 15 S. 3 DSM-RL benennen kann: *Heesen/Jüngels*, RuZ 2021, 45, 52.

⁹⁹ Gesetzentwurf der Bundesregierung, Entwurf eines Gesetzes zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes, BT-Drs. 19/27426, S. 100.

¹⁰⁰ *Raue*, ZUM 2021, 793, 798.

¹⁰¹ Ausführlich hierzu *Raue*, ZUM 2021, 793, 801: Der Rechteinhaber muss nach § 95b Abs. 1 Nr. 11 UrhG den Begünstigten die notwendigen Mittel zur Verfügung zu stellen, damit sie von der TDM-Schranke trotz technischer Schutzmaßnahmen in dem erforderlichen Maße Gebrauch machen können.

¹⁰² *Dreier*, in *Dreier/Schulze*, Urheberrechtsgesetz, 7. Aufl. 2022, § 60g Rn. 3.

¹⁰³ *Raue*, ZUM 2021, 793, 798.

Einzel Forscher (Abs. 3). Öffentlich zugänglich¹⁰⁴ gemacht werden dürfen die Vervielfältigungen einem abgegrenzten Personenkreis für die gemeinsame wissenschaftliche Forschung sowie einzelnen Dritten zur Qualitätsüberprüfung. Eine Präsentation der Vervielfältigungen über diese Grenzen hinaus wird durch die TDM-Schranke nicht ermöglicht. Berechtigte – nicht hingegen Einzel Forscher – dürfen nach § 60d Abs. 5 UrhG die Vervielfältigungen mit angemessenen Sicherheitsvorkehrungen gegen unbefugte Benutzung aufbewahren, solange sie für die Zwecke der wissenschaftlichen Forschung oder zur Überprüfung der wissenschaftlichen Qualität erforderlich sind. Auch das Text und Data Mining für die wissenschaftliche Forschung wird nach § 60h Absatz 2 Nr. 2 UrhG vergütungsfrei gestellt.

Wer Vervielfältigungsstücke über die beschriebenen Freistellungen der TDM-Schranken hinaus nutzen oder präsentieren möchte, muss die Zustimmung der Rechteinhaber einholen oder sich auf eine der übrigen Schranken der §§ 44a ff. UrhG berufen können. Abgeleitete Textformate können einen Ansatz für Lösungen des urheberrechtskonformen Zugangs zu großen Textkorpora außerhalb des Rahmens der Text und Data-Mining Schranken bieten.¹⁰⁵

Leseempfehlungen zur vertiefenden Lektüre: *Raue*, Die Freistellung von Datenanalysen durch die neuen Text und Data Mining-Schranken (§§ 44b, 60d UrhG), *Zeitschrift für Urheber- und Medienrecht (ZUM)* 2021, 793, *Heesen/Jüngels*, Der Regierungsentwurf der Text und Data Mining-Schranken (§§44b und 60d UrhG-E), *Recht und Zugang (RuZ)* 2021, 45, *Spindler*, Die neue Urheberrechts-Richtlinie der EU, insbesondere „Upload-Filter“ – Bittersweet?, *Computer und Recht (CR)* 2019, 277, *Raue/Schöch*, Zugang zu großen Textkorpora des 20. und 21. Jahrhunderts mit Hilfe abgeleiteter Textformate – Versöhnung von Urheberrecht und textbasierter Forschung, *Recht und Zugang (RuZ)* 2020, 118.

VI. Literaturverzeichnis

- David M. Blei*, Probabilistic Topic Models, <http://www.cs.columbia.edu/~blei/papers/Blei2012.pdf> (abgerufen am 20.07.2021).
- Kai-Uwe Carstensen, Christian Ebert, Cornelia Ebert, Susanne Jekat, Ralf Klabunde, Hagen Langer (Hrsg.)*, Computerlinguistik und Sprachtechnologie, 3. Aufl. 2010, Spektrum Akademischer Verlag, Heidelberg.
- Katharina de la Durantaye, Linda Kuschel*, Regelungen zu nicht verfügbaren Werken, *Zeitschrift für Urheber- und Medienrecht (ZUM)* 2021, 785.
- Thomas Dreier, Gernot Schulze (Hrsg.)*, Urheberrechtsgesetz, 7. Aufl. 2022, C. H. Beck München.
- Thomas Dreier, Gernot Schulze (Hrsg.)*, Urheberrechtsgesetz, 6. Aufl. 2018, C. H. Beck München.
- Katharina Erler-Fridgen*, Die Nutzung wissenschaftlicher Ausgaben für Textanalysen, *IRDT PAPERSERIES Nr. 1*.

¹⁰⁴ Zum Begriff des „Öffentlich Zugänglichmachens“ in § 60d Abs. 4 UrhG siehe oben.

¹⁰⁵ *Raue/Schöch*, *RuZ* 2020, 118, 120.

- Katharina Erler-Fridgen*, Kriterien der urheberrechtlichen Schutzfähigkeit von Texten und Sammelwerken, IRDT PAPERSERIES Nr. 2.
- Katharina Erler-Fridgen*, Die Präsentation von Textteilen als Ergänzung von Textanalysen, IRDT PAPERSERIES Nr. 3.
- Katharina Erler-Fridgen*, Datenbanken als Quelle oder Ergebnis von Textanalysen – Datenbankwerkschutz und das Leistungsschutzrecht des Datenbankherstellers, IRDT PAPERSERIES Nr. 4.
- Hendrik Heesen, Leonie Jüngels*, Der Regierungsentwurf der Text und Data Mining-Schranken (§§44b und 60d UrhG-E), Recht und Zugang (RuZ) 2021, 45.
- Gabe Ignatow, Rada Mihalcea*, Text Mining – A Guidebook for the Social Sciences, 2017, Sage Publications Thousand Oaks/London/New Dehli/Singapore.
- Felicitas Kleinkopf, Janina Jacke, Markus Gärtner*, Text- und Data-Mining, Urheberrechtliche Grenzen der Nachnutzung wissenschaftlicher Korpora bei computergestützten Verfahren und digitalen Ressourcen, Multimedia und Recht (MMR) 2021, 196.
- Fotis Jannidis, Hubertus Kohle, Malte Rehbein (Hrsg.)*, Digital Humanities – Eine Einführung, 2017, J.B. Metzler Verlag, Stuttgart.
- Ulrich Loewenheim, Matthias Leistner, Ansgar Ohly (Hrsg.)*, Schricker/Loewenheim, Urheberrecht, 6. Auflage 2020, C.H. Beck München.
- Benjamin Raue*, Text und Data Mining, Die neue Urheberrechtsschranke des § 60d UrhG, Computer und Recht (CR) 2017, 656.
- Benjamin Raue*, Das Urheberrecht der digitalen Wissen(schaft)sgesellschaft, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2017, 11.
- Benjamin Raue*, Free Flow of Data? The Friction Between the Commission’s European Data Economy Initiative and the Proposed Directive on Copyright in the Digital Single Market, International Review of Intellectual Property and Competition Law (IIC) 2018, 379.
- Benjamin Raue*, Rechtssicherheit für datengestützte Forschung, Die Text-und-Data-Mining-Schranken in Art. 3 und 4 DSM-RL, Zeitschrift für Urheber- und Medienrecht (ZUM) 2019, 684.
- Benjamin Raue*, Die geplanten Text und Data Mining-Schranken (§§ 44b und 60d UrhG-E), Zeitschrift für Urheber- und Medienrecht (ZUM) 2020, 172.
- Benjamin Raue*, Die Freistellung von Datenanalysen durch die neuen Text und Data Mining-Schranken (§§ 44b, 60d UrhG), Zeitschrift für Urheber- und Medienrecht (ZUM) 2021, 793.
- Benjamin Raue, Christof Schöch*, Zugang zu großen Textkorpora des 20. und 21. Jahrhunderts mit Hilfe abgeleiteter Textformate – Versöhnung von Urheberrecht und textbasierter Forschung, Recht und Zugang (RuZ) 2020, 118.
- Haimo Schack*, Schutzgegenstand, „Ausnahmen oder Beschränkungen“ des Urheberrechts, Gewerblicher Rechtsschutz und Urheberrecht (GRUR) 2021, 904.

Gerald Spindler, Die neue Urheberrechts-Richtlinie der EU, insbesondere „Upload-Filter“ – Bittersweet?, Computer und Recht (CR) 2019, 277.