

Digital Humanities und Recht

Zu den neuen Regeln für das Text und Data Mining (TDM) und ihrem strategischen Potential für die Bibliotheken

Frédéric Döhl und Dorothea Zechmann

Die Digital Humanities sind ein ebenso dynamischer wie heterogener Bereich, mit einem komplexen Diskurs um ihre Möglichkeiten und Grenzen, ihre Identität als akademischer Disziplin und ihrem Verhältnis zu traditionellen geisteswissenschaftlichen Fächern und Forschungsstrategien.¹ Die Digital Humanities befinden sich dabei derzeit in vielfacher Hinsicht in einer Experimentierphase, nicht nur hinsichtlich ihrer institutionellen Formate, Verortung und Verstetigung, sondern auch hinsichtlich der sie leitenden wissenschaftlichen Theorien, Methoden und Ziele.²

Einiges lässt sich jedoch als prägend für den Bereich herausdestillieren, ohne tiefer in diesen weithin ausdifferenzierten, ebenso projekt- wie publikationsreichen Diskurs eintreten zu müssen. Dazu zählt u. a., dass Digital Humanities zuvorderst geisteswissenschaftliche Arbeit mit substantieller Computerunterstützung betreibt und bezeichnet, regelmäßig also Forschung an digital vorliegenden Informationen. Führende Vertreter der Digital Humanities sprechen insoweit auch von „rechnenden Verfahren“ (Gerhard Lauer), „computergestützten Methoden“ (Fotis Jannidis) oder „rechnerabhängigen Verfahren“ (Sybille Krämer/Martin Huber), von „digitalen Verfahren“ (Patrick Sahle) oder „digitalen Methoden“ (Claudine Moulin, Caroline Sporleder).³ Als Kernbereiche der Digital Humanities haben sich dabei die digitale Editi-

Der Beitrag setzt sich mit der anstehenden Urheberrechtsreform im Bereich des Text und Data Mining und der Frage auseinander, was sich hierdurch für die Bibliotheken voraussichtlich ab Juni 2021 ändern wird hinsichtlich Möglichkeiten, Erwartungen und Verantwortung.

onphilologie, die digitale Ergebnispräsentation (Visualisierung, Rekonstruktion) und das Text und Data Mining (TDM) herausgebildet. Um letzteres geht es hier.

Text und Data Mining, Bibliotheken und Recht

Text und Data Mining ist der Oberbegriff „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“⁴. Daten schließt dabei Medien jedweder Art ein, also z. B. auch Töne, Bilder und Bewegtbilder. Bei TDM geht es um die computerunterstützte geisteswissenschaftliche Forschung an und mit großen, händisch nicht mehr mit vertretbarem Aufwand oder schlicht gar nicht mehr händisch bewältigbaren digitalen bzw.

1 Vgl. für eine Kurzeinführung Frédéric Döhl: „Was sind Digital Humanities?“, in: *Dialog mit Bibliotheken* 32/1 (2020), S. 18-24, <<https://d-nb.info/1206109068/34>>, sowie ausführliche mit vielen weiteren Nachweisen Frédéric Döhl: „Digital Humanities und Bibliotheken. Über technisch-organisatorische Infrastruktur hinausgedacht“, in: *ZfBB – Zeitschrift für Bibliothekswesen und Bibliographie* 66/1 (2019), S. 4-18; Frédéric Döhl: „Digital Turn – Gedächtnisinstitutionen und Digital Humanities. Zwischenbericht aus seiner Workshop-Reihe der Deutschen Nationalbibliotheken“, in: *ZfBB – Zeitschrift für Bibliothekswesen und Bibliographie* 67/3-4 (2020), S. 213-230.

2 Vgl. Fotis Jannidis: „Digitale Geisteswissenschaften: Offene Fragen – schöne Aussichten“, in: *Zeitschrift für Medien- und Kulturforschung* 10/1 (2019), S. 63-70; Gerhard Lauer: „Über den Wert der exakten Geisteswissenschaften“, in: *Geisteswissenschaft – was bleibt? Zwischen Theorie, Tradition und Transformation*, hrsg. von Hans Joas/Jörg Noller, Verlag Karl Alber: Freiburg im Breisgau 2019, S. 152-173.

3 Gerhard Lauer: „Die digitale Vermessung der Kultur. Geisteswissenschaften als Digital Humanities“, in: *Big Data. Das neue Versprechen der Allwissenheit*, hrsg. von Heinrich Geiselberger, Suhrkamp: Berlin 2013, S. 99-116, hier S. 101; Fotis Jannidis: „Digitale Geisteswissenschaften: Offene Fragen – schöne Aussichten“, in: *Zeitschrift für Medien- und Kulturforschung* 10/1 (2019), S. 63-70, hier S. 69; Sybille Krämer/Martin Huber: „Dimensionen Digitaler Geisteswissenschaft. Zur Einführung in diesen Band“, in: *Wie Digitalität die Geisteswissenschaften verändert: Neue Forschungsgegenstände und Methoden*, hrsg. von dens., 2018, <www.zfdg.de/sb003_013>; Patrick Sahle: „Digital Humanities? Gibt's doch gar nicht!“, in: *Zeitschrift für digitale Geisteswissenschaften* (2015), <http://www.zfdg.de/sb001_004>; Claudine Moulin, zitiert nach Christian Möller: „Digital Humanities. Wie die Digitalisierung die Wissenschaft verändert“, in: *DLF* (5.3.2015), <https://www.deutschlandfunkkultur.de/digital-humanities-wie-die-digitalisierung-die-wissenschaft.976.de.html?dram:article_id=313420>; Caroline Sporleder: Was sind eigentlich Digital Humanities?, 2013, <<https://www.academics.de/ratgeber/digital-humanities-berufschancen>>.

4 Art. 2 Nr. 2 Richtlinie (EU) 2019/790 des Europäischen Parlaments 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, 2019, <<https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>>.

digitalisierten Informationsbeständen. Das ist in der Krebsforschung, der Weltraumforschung oder der Klimaforschung Standard. In den Geisteswissenschaften, die sich in vielen Bereichen für relativ überschaubare, oft zudem sehr gut erschlossene Gruppen von Werken, Quellen und Praktiken interessieren und in denen zugleich das Seltene, Besondere und Ungewöhnliche oft besonders im Fokus steht, ist diese dezidierte Big-Data-Perspektive zwar nicht neu – man denke nur an die Nutzung statistischer Methoden empirischer Sozialforschung –, aber in dieser disziplinären Breite und mit diesem politischen und infrastrukturellen Nachdruck wie derzeit zu beobachten doch von einer neuen Qualität.

Das ist wissenschaftspolitisch eine durchaus ambivalente, in Teilen immer wieder auch umstrittene Herausforderung. Und wie weit diese ›Digitalisierung der Geisteswissenschaften‹ am Ende gehen wird, ist in der Tat derzeit schwer abzuschätzen. Nicht zuletzt, weil sich viele geisteswissenschaftlich relevante Forschungsgegenstände, die auch nichts an ihrer Relevanz einbüßen, schlecht und mitunter auch gar nicht in digital auswertbare Daten übersetzen lassen. Man denke nur an performative Praktiken oder an ästhetische Erfahrungen, an physische Objekte oder an zentrale historische Ereignisse wie entscheidende politische Verhandlungen, anlässlich deren schlicht nur zwei, drei Leute im Raum waren, und an deren mutmaßlichen Verlauf man sich nur indirekt über eine Folgenanalyse annähern kann. Ebenso an Zivilisationen, die bis auf wenige Zeugnisse schlicht verschwunden sind.

Aber datenbasierte Forschung ist eben in bestimmten geisteswissenschaftlichen Bereichen über alle Fachdisziplinen hinweg sehr wohl möglich, insbesondere da, wo der jeweilige Forschungsgegenstand in seiner Primärquelle bereits selbst als digitales oder unmittelbar digitalisierbares Datenformat vorliegt – Buchtext oder Musik-CD z.B. – oder man über digitale bzw. digitalisierbare Sekundärquellen zumindest eine substantielle Annäherung an ihn schaffen kann, z.B. über Photographien von Deckengemälden oder Transkriptionen von Jazzsoli, um nur zwei Beispiele aus der aktuellen Digital-Humanities-Forschung zu nennen.⁵ Und insoweit ist eben klar, dass die Geisteswissenschaften in ihrer traditionellen Manier an vielen Stellen schon notgedrungen bestimmte Frage-

stellungen gar nicht angehen konnten, weil sich Informationsbestände ab einer gewissen Größe gar nicht befragen ließen. Schon die knapp sieben Regalkilometer deutschsprachiger Buch- und Musikproduktion mit allein gut 70.000 physischen Musikträgern, die nach wie vor alljährlich die Regale der Deutschen Nationalbibliothek erreichen, kann niemand mehr lesen bzw. hören.⁶ Von der Fähigkeit zum Vergleich etwa über viele Sprachgrenzen hinweg ganz zu schweigen. Bevor man aber etwas ausprobiert, kann man sinnvollerweise gar nicht bemessen, welchen Wert und Ertrag es hat.

Vor diesem Hintergrund interessiert Digital Humanities u.a., kanonisiertes Wissen in größeren Informationskontexten zu überprüfen, zu vernetzen und wo möglich durch Annotation und Metadaten anzureichern, aber auch in Relation zu setzen zu einem neuen Wissen über jenes Gros kultureller Produktion, über das man allenfalls oberflächliche Kenntnisse hat und das schlicht nicht, geschweige denn in gleicher Weise präsent ist wie jene Produktionen und Praktiken, die Gegenstand kanonisierten Wissens sind – man denke an Alltagskultur und Kleinmeister, Vergessenes und aus der Mode Gekommenes.⁷ Und nicht zu vergessen ist schließlich die Gegenwart selbst als Forschungsfeld, in der sich vieles schlicht digital vollzieht und digitaler Methoden bedarf.

Text und Data Mining ist der Bereich, in dem vor allem hieran gearbeitet wird. Hier können die *Bibliotheken* einen wertvollen Beitrag leisten. Man denke nur an den Bestand der Deutschen Nationalbibliothek mit gut 40 Millionen Medieneinheiten, die mit deren Sammlung und Erschließung aufgrund des gesetzlichen Auftrags mit viel Aufwand darauf zielt, Bestandteile der Sammlung auffindbar und nutzbar zu halten. Mit Text und Data Mining kann es u. a. gelingen, eine solche Sammlung und deren Inhalte als Wissensspeicher auch in größeren Zusammenhängen zu verstehen und für den Bibliotheksnutzer produktiv zu machen.

Nur wenige Bibliotheken mit regelmäßig sehr wissenschaftsnahem Aufgabenprofil treten freilich bislang in diesem Digital-Humanities-Kontext verstärkt in Erscheinung, wie man stellvertretend z.B. daran ablesen kann, wer wiederkehrend aktiv mit Input hierzu an den seit 2014 stattfindenden Jahrestagungen des zentralen DH-Fachverbandes DHD (Digital Humanities im deutschsprachigen Raum) teilnimmt.⁸ Und selbst

5 Vgl. <<https://deckenmalerei.badw.de/aktuelles.html>>; <<http://dig-that-lick.eecs.qmul.ac.uk>>.

6 Vgl. Deutsche Nationalbibliothek: Jahresbericht 2018, <<https://d-nb.info/1187664855/34>>, S. 35, 44.

7 Fotis Jannidis: Digitale Literaturwissenschaft. In: Digitalität. Theorien und Praktiken des Digitalen in den Geisteswissenschaften, 2016. Verfügbar unter: <https://diggeist.hypothesos.org/114>.

8 Vgl. stell. die Abstracts der DHD-Jahrestagungen 2015 bis 2020 <<http://gams.uni-graz.at/o:dhd2015.abstracts-gesamt>>, <http://dhd2016.de/sites/default/files/dhd2016/files/boa-2.0_ohne_Vorwort.pdf>, <http://www.dhd2017.ch/wp-content/uploads/2017/02/Abstractband_ergaenzt.pdf>.

jene Bibliotheken, die schon einen verstärkten Bezug zu den Digital Humanities, bisweilen sogar eigene proaktive Strukturen von Planstellen bis Digital-Humanities-Labs ausgebildet haben, sind bislang kaum in diesem Kernbereich der Digital Humanities engagiert, nämlich der besagten Analyse und Interpretation von Daten zur Beantwortung geisteswissenschaftlicher Fragen bzw. Überprüfung geisteswissenschaftlicher Thesen einschließlich der Entwicklung dieser Fragen/Thesen, wie der aktuelle, 2019 erarbeitete Report „Europe’s Digital Humanities Landscape“ der LIBER Digital Humanities & Digital Cultural Heritage Working Group zeigt.⁹ Die Domäne der Bibliotheken in Digital-Humanities-Projekten liegt bislang im Bereich von Vor-, Zu- und Nacharbeiten auf der ganzen Skala bibliothekarischer Serviceangebote, ihre Rolle wird auch von den Forschenden entsprechend verortet.¹⁰ Freilich ist diese klassische Rollenverteilung derzeit in Bewegung. Darauf wies bereits jüngst die PARTHENOS-Studie zur mittelfristigen Weiterentwicklung der Digital Humanities hin, 2019 im Rahmen des Horizon-2020-Programms der EU-Kommission erstellt. Hier werden die Wissenschaft und Gedächtnisinstitutionen nicht hierarchisch, sondern ausdrücklich auf Augenhöhe adressiert: „The [PARTHENOS] foresight study was designed to address how digital research methods in the digital humanities and cultural heritage sector may develop over the next 5-10 years [...]“¹¹ Eine Perspektive, die ebenfalls bereits angelegt ist in den geisteswissenschaftlich orientierten und dabei kooperativ organisierten Absichtserklärungen, Anträgen und Gutachten im Rahmen des laufenden Prozesses zum Aufbau einer Nationalen Forschungsdateninfrastruktur (NFDI), der gerade überall Wissenschaft und Gedächtnisinstitutionen in engeren Austausch und Teamarbeit an gemeinsamen digitalen Herausforderungen bringt und dies im Fall positiver Bescheide für

geisteswissenschaftlich orientierte NFDI-Konsortien nochmals deutlich forcieren dürfte.¹²

Diese Entwicklung kommt nicht von ungefähr, so in den Digital Humanities selbst immer wieder auf eine sich derzeit verändernde Rolle der Bibliotheken für die und in den Digital Humanities hingewiesen wird, etwa jüngst von Gerhard Lauer: „Wir hingegen wollen über Digital Humanities die Bibliothek viel dichter mit den Fächern verzahnen, ‚verzapfen‘ würde man schreinerisch sagen. In der Vorstellung vieler Geisteswissenschaftlerinnen und -wissenschaftler ist die Bibliothek ein Ort, an dem man ein Buch entleiht, liest oder zur Anschaffung bestellt. Dass Bibliotheken schon länger viel mehr sind, ist wenig bekannt. Die Bewegung der Digital Humanities könnte nun mit sich bringen, dass wir die Fächer wieder besser mit der Bibliothek verzahnen oder sogar verzapfen.“¹³ Das anstehende neue Recht verspricht an dieser Stelle eine nochmals potenzierte Dynamik ganz in Lauers Sinne, hin zu einem intensivierten Zusammenspiel von Forschung und Bibliotheken.

„Angesichts der Entwicklungsdynamik digitaler Systeme und Arbeitsmöglichkeiten ist – beispielsweise in den Digital Humanities – die infrastrukturelle Unterstützung des gesamten wissenschaftlichen Wertschöpfungsprozesses sowie dessen Öffnung über den akademischen Kreis hinaus gefordert“¹⁴, mahnten Theo Hug und Petra Missomelius schon 2015 an. Der von ihnen benannte Blick auf den gesamten wissenschaftlichen Wertschöpfungsprozess ist es nun, der an dieser Stelle jetzt in den Fokus rückt und eine nochmals neue Perspektive erhält. Er tut dies in zweifacher Hinsicht. Und zwar im Teilbereich des Text und Data Mining zum Zwecke wissenschaftlicher Forschung. Hintergrund ist die im Frühjahr 2019 beschlossene *Richtlinie über das Urheberrecht im digitalen Binnenmarkt 2019/790 der Europäischen Union* (kurz:

<https://zenodo.org/record/3684897#.XnngS2X-u4>, <https://zenodo.org/record/2596095#.Xnngli2X-u4>, <https://zenodo.org/record/3666690#.XnngQi2X-u4>.

- 9 Lotte Wilms/Caleb Derven/Liam O’Dwyer/Kirsty Lingstadt/Demmy Verbeke: Europe’s Digital Humanities Landscape: A Study from LIBER’s Digital Humanities & Digital Cultural Heritage Working Group, 2019, S. 15, https://zenodo.org/record/3247286#.XnoDOy2X_q1.
- 10 DHd: Digital Humanities 2020, 2014, <https://www.dhd2014.uni-passau.de/fileadmin/dokumente/projekte/digitalhumanities/DH2020-2014-2-Lang.pdf>; Ben Kaden: „Zur Epistemologie digitaler Methoden in den Geisteswissenschaften“, in: Berliner Beiträge zu Digital Humanities, 2016, S. 2, Preprint: <https://zenodo.org/record/50623#.XocCBC35yCU>; Rafael Ball: „Die Bibliothek der Zukunft [Editorial]“, in: b.i.t.online 22/3 (2019), S. 201, <https://www.b-i-t-online.de/heft/2019-03-editorial.pdf>; Mareike König: Was sich Historiker/-innen von Archiven wünschen: eine Umfrage, 2019, <https://dhdhi.hypotheses.org/6107>.
- 11 Vgl. Mark Hedges et. al.: Digital Humanities Foresight, The Future Impact of Digital Methods, Technologies, and Infrastructures (= DARIAH-DE Working Papers Nr. 40), 2019, S. 7, <http://webdoc.sub.gwdg.de/pub/mon/dariah-de/dwp-2019-40.pdf>. Für die volle PARTHENOS foresight study vgl. https://zenodo.org/record/2662490#.Xh8RrS2X_dc. Für eine Bewertung aus Bibliothekssicht vgl. Ulrike Wuttke/Claus Spiecker/Heike Neuroth: „PARTHENOS – Eine digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften“, in: Bibliothek – Forschung und Praxis 43/1 (2019), S. 11-20.
- 12 Vgl. nur die Absichtserklärungen der geplanten Konsortien unter <https://www.dfg.de/foerderung/programme/nfdi/absichtserklaerungen/index.html>.
- 13 Gerhard Lauer in Andreas Brandtner/Gerhard Lauer/Peter Reuter: „Die Bibliotheken haben ihre Zukunft vor sich, aber es sind Bibliotheken des 21. Jahrhunderts.“ Bibliotheken als Infrastrukturen der Geisteswissenschaften und als Orte der Selbstkultivierung: Ein Gespräch mit Gerhard Lauer von Andreas Brandtner und Peter Reuter“, in: ABI Technik 39/2 (2019), S. 171-178, hier S. 175.
- 14 Theo Hug/Petra Missomelius: „Universitätsbibliotheken, digitale Medien und Mobilität“, in: Die Bibliothek in der Zukunft. Regional – Global: Lesen, Studieren und Forschen im Wandel. Festschrift für Hofrat Dr. Martin Wieser anlässlich seiner Versetzung in den Ruhestand, hrsg. von Klaus Niedermair/Dietmar Schuler, Innsbruck University Press: Innsbruck 2015, S. 203-222, hier S. 206.

DSM-RL).¹⁵ Diese ist nun von den EU-Mitgliedsstaaten bis Juni 2021 in nationales Recht umzusetzen. Nach einem Stellungnahmeverfahren im Herbst 2019 wurde im Januar 2020 ein Diskussionsentwurf des Bundesministeriums für Justiz und Verbraucherschutz (BMJV) für die Umsetzung von ersten Teilen des Richtlinieninhalts veröffentlicht und kommentiert. Derzeit befindet sich der Referentenentwurf in der Ressortabstimmung des Bundes.

Die DSM-RL reguliert auch einen zentralen Bereich der Digital Humanities neu und EU-einheitlich: Text und Data Mining (Art. 2-4) – genauer: die Modalitäten, wann und in welcher Form Text und Data Mining jedenfalls urheber- und leistungsschutzrechtlich zulässig ist

Die Zulässigkeit von Text und Data Mining ist deswegen entscheidend für diesen Forschungsbereich, weil es hierbei regelmäßig um Massenverfahren mit großen Datenmengen geht, die Einzellizensierungen, ja selbst schon wissenschaftliche Quellenangaben im Sinne des Zitatrechts regelmäßig faktisch unmöglich machen. Und sie ist aus Sicht des EU-Gesetzgebers als Beschränkung der betroffenen Urheber- und Leistungsschutzrechte angemessen, da es bei Text und Data Mining nicht um Werkgenuss im urheberrechtlich geschützten Sinne geht, sondern ‚lediglich‘ um Informationen in den Werken, die als Informationen im Übrigen meist für sich genommen ohnehin nicht schutzfähig sind. Im Übrigen kommt es weder zu einer Beeinträchtigung der Persönlichkeitsrechte des Urhebers noch der Verwertungsmöglichkeiten von Urhebern und Leistungsschutzberechtigten am Werk und seinen medialen Manifestationen als Buch, Tonträger, Film usw.¹⁶

Fehlt eine rechtssichere Arbeitsgrundlage dieser Art, bleibt der Digital-Humanities-Forschung nur die Arbeit mit gemeinfreien Materialien. Das hat weitreichende Konsequenzen. In der Gegenwart bleibt daher nur ein Bruchteil an Forschungsmaterial. Dieser Umstand hat ohne wissenschaftliche Gründe zu einem Forschungsschwerpunkt der Digital Humanities im sogenannten „langen 19. Jahrhundert“ (bis etwa zum Ende des Ersten Weltkriegs) geführt hat. Befördert dadurch, dass parallel zu den rechtlichen Gegebenheiten noch andere bestehen (weiter zurück als ca. 1800 ist z. B. die

Qualität der Resultate automatischer Texterkennung/OCR regelmäßig [noch] nicht gut genug für derartige Forschungsarbeit). Neben Konsequenzen für den zeitlichen Forschungsschwerpunkt hatte das bisherige Recht zugleich zudem Auswirkungen auf die Priorisierung der bevorzugten Forschungsmedien. Faktisch wurde hierdurch die Konzentration auf das Medium Text nachhaltig befördert, was sofort klar wird, wenn man sich erinnert, wie alt der Buchdruck ist (ab ca. 1450 – Gutenberg), wie spät aber zugleich andere für unsere Gegenwart zentrale Medien überhaupt erst erfunden wurden – etwa Schwarzweißfotografie (ab ca. 1839 – Louis Daguerres Daguerreotypie), Farbfotografie (ab ca. 1861 – James Clerk Maxwells erstes dauerhaftes Farbbild), Tonaufnahme (ab ca. 1877 – Thomas Alva Edisons Phonograph) und Film (ab ca. 1888 – Louis Le Princes Roundhay Garden Scene). Und vom Erreichen qualitativer hochwertiger Marktreife dieser Medien ist da noch gar nichts gesagt; Tonaufnahmen z.B. beginnen durch eine Reihe damaliger Innovationen in der Produktionstechnologie überhaupt erst mit den 1950er Jahren, für heutige Ohren nicht mehr historisch zu klingen. Aufgrund der Schutzfristen steht jedoch nur ein Bruchteil dieser kulturellen Produktion der Forschung bislang für wissenschaftliche Zwecke zur Nutzung frei zur Verfügung.

Soll also Forschung im Zeithorizont des „living memory“ in seiner ganzen medialen Vielfalt möglich sein, bedarf es einer solchen Beschränkung von Rechten zugunsten einer zulässigen Nutzung von Werken im Wege von Text und Data Mining unabhängig davon, ob diese Werke bereits gemeinfrei oder urheberrechtlich geschützt sind. Dies ist nun vom EU-Gesetzgeber gewollt. Denn „Geisteswissenschaften, die diese Daten und Informationen nur wenigen Internetfirmen überlassen, hören auf, sich ein Wissen von der Gesellschaft in Wissenschaftsform zu verschaffen“¹⁷, wie Gerhard Lauer anmerkt.

Um dieses Ziel zu fördern, heißt es nun in Art. 3 Abs. 1 der DSM-RL: „Die Mitgliedsstaaten 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

15 Vgl. Richtlinie (EU) 2019/790 des Europäischen Parlaments 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. Verfügbar unter <<https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>>. Vgl. zum Diskussionsstand z. B. die Stellungnahmen im laufenden Umsetzungsverfahren des BMJV, <https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/DE/Gesetz_I_Umsetzung_EU_Richtlinien_Urheberrecht.html>.

16 Vgl. Erwägungsgrund 16 der Richtlinie (EU) 2019/790 des Europäischen Parlaments 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. Verfügbar unter <<https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>>.

17 Gerhard Lauer: „Über den Wert der exakten Geisteswissenschaften“; in: Geisteswissenschaft – was bleibt? Zwischen Theorie, Tradition und Transformation, hrsg. von Hans Joas/Jörg Noller, Verlag Karl Alber: Freiburg im Breisgau 2019, S. 152-173, hier S. 170.

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.“¹⁸ Die technischen Verweise zu Beginn dieser Norm benennen dabei all jene Rechte, die durch diesen neuen Art. 3 Abs. 1 DSM-RL künftig zugunsten eines zulässigen wissenschaftlichen Text und Data Mining beschränkt werden, etwa das Urheberrecht (Art. 2 Buchst. a InfoSoc-RL 2001/29/EG) oder das Presseverlegerleistungsschutzrecht (Art. 15 Abs. 1 DSM-RL).

Wissenschaftliche Forschung, wie es als Bedingung begrifflich in Art. 3 Abs. 1 der DSM-RL heißt, ist hier nicht näher definiert, meint aber im Sinne des Art. 13 EU-Grundrechtecharta „jede methodische und systematische Tätigkeit“, „die das Ziel hat, in nachprüfbarer Weise neue Erkenntnisse zu gewinnen“¹⁹. Weitere Qualifikationen, geschweige denn formale Anerkennungsprozesse (etwa durch die Deutsche Forschungsgemeinschaft) sind an dieser Stelle nicht zu erfüllen, außer einem Prozessmanagement, das eben diese Transparenz und Nachprüfbarkeit sicherstellt.

Das anstehende neue Recht des Text und Data Mining

Ausgangspunkt unseres Beitrags ist nun die Feststellung, dass sich das neue Recht des Text und Data Mining eben keineswegs mehr nur an Forschende im klassischen Sinne allein richtet. D. h. insbesondere nicht nur an Angehörige von Universitäten und Hochschulen sowie von außeruniversitären öffentlich-rechtlichen Forschungseinrichtungen (wie Fraunhofer-Gesellschaft, Helmholtz-Zentren, Leibniz-Gemeinschaft, Max-Planck-Gesellschaft und Akademien der Wissenschaften). Neben diese treten in einer partnerschaftlichen Rolle ausdrücklich die Kulturerbeeinrichtungen. Gemeint sind damit „öffentlich zugängliche Bibliotheken, Archive, Einrichtungen des Film- und Tonerbes und öffentlich zugängliche Museen“²⁰ (so die Formulierung im Diskussionsentwurf des BMJV vom 15. Januar 2020 zur Umsetzung von Art. 3 DSM-RL). Dabei kommt diesen privilegierten Institutionen eine besondere Verantwortung zu.

Sie sind nun aktive Akteure und Adressaten, ermöglichen nicht nur Text und Data Mining zu Forschungszwecken, sondern sie haben auch die Verpflichtung, die so gewonnenen Informationen gegen unbefugte Benutzung mit angemessenen Sicherheitsvorkehrungen zu schützen (siehe Entwurf § 60d Abs. 5 UrhG-E). Insofern sind sie neben ihrem tätigen Part auch Mittler und Moderatoren zwischen Rechteinhabern einerseits und Wissenschaftlern andererseits. Die Kulturerbeeinrichtungen werden nun gleichberechtigt als aktive Akteure eines solchen Text und Data Mining zum Zwecke wissenschaftlicher Forschung adressiert und privilegiert. Und – das ist wichtig zu verstehen – grundsätzlich unabhängig von den anderen privilegierten Akteursgruppen, d. h. von Forschungsorganisationen und Einzelforschern.

Mit dieser Entscheidung des EU-Gesetzgebers scheint also mehr intendiert zu sein als eben jene klassischen bibliothekarischen Vor-, Zu- und Nacharbeiten zur Unterstützung von externen Forschenden bei der Durchführung von Text und Data Mining an und mit den eigenen Beständen. Die Erwägungsgründe der Richtlinie konkretisieren die Intention des neuen Rechts freilich nicht näher. Es wird lediglich festgestellt, dass man auch Einrichtungen des Kulturerbes mit einbeziehen wolle, „da diese möglicherweise ebenfalls Forschung im Zusammenhang mit ihrer hauptsächlichen Tätigkeit betreiben könnten.“²¹ Der Begleittext zum derzeit zirkulierenden deutschen Umsetzungsentwurf von Art. 3 DSM-RL in nationales Recht benennt auch nur vergleichsweise vage und allgemein gehalten Innovationsförderung als Motivation des Gesetzgebers. Gemeint sind damit aber Existenzrecht und Legitimation einer solchen Schranke zugunsten einer zulässigen Nutzung von Werken zu Text und Data Mining zum Zwecke wissenschaftlicher Forschung insgesamt, nicht gesondert die Einbeziehung der Einrichtungen des Kulturerbes hierin. Faktisch also sehen sich die Bibliotheken absehbar einem neuen Recht gegenüber, dass sie direkt anspricht und ihnen ausdrücklich etwas anbietet, ohne aber einen Beipackzettel beigefügt bekommen zu haben, was denn nun damit wie angefangen werden soll. Das anstehende neue Recht lässt an dieser Stelle also nicht nur eine nochmals potenzierte Dynamik erwarten, hin zu einem intensivierten Zusammenspiel von Forschung

18 Art. 3 Abs. 1 Richtlinie (EU) 2019/790 des Europäischen Parlaments 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, 2019, <<https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>>.

19 Thomas Dreier/Gernot Schulze: Urheberrechtsgesetz, 6. Aufl., C.H. Beck: München 2018, § 60d UrhG, Rn 5.

20 Vgl. BMJV, <https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/DE/Gesetz_I_Umsetzung_EU_Richtlinien_Urheberrecht.html>.

21 Erwägungsgrund 8 der Richtlinie (EU) 2019/790 des Europäischen Parlaments 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, 2019, <<https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>>.

und Bibliotheken im Bereich Text und Data Mining. Es bringt auch Klärungsbedarf für das Bibliothekswesen. Wir möchten mit unserem Beitrag gerne einen entsprechenden Diskurs in diese Richtung befördern, so wie es Anliegen des ursprünglich geplanten Vortrags auf dem Bibliothekartag in Hannover gewesen war. Denn auch wir sehen uns als Deutsche Nationalbibliothek mit dieser eben zugleich Chance wie Herausforderung konfrontiert, die der Gesetzgeber nun an uns heranträgt.

Wichtig zu verstehen ist, dass sich keineswegs alles ändert mit Umsetzung der DSM-RL in nationales Recht. Das Gegenteil ist der Fall. Das bis spätestens Juni 2021 national umzusetzende neue Recht wird vor allem viel davon bestätigen und EU-weit angleichen, was in Deutschland durch das UrhWissG – anders als in vielen anderen EU-Mitgliedsstaaten – ohnehin schon seit März 2018 Recht ist und damit derzeit Arbeitsgrundlage aller Bibliotheken. Die aktuell geltende Regelung zum Text und Data Mining findet sich in § 60d des Urheberrechtsgesetzes (UrhG).²² Deutschland hatte in dieser Sache auf die an anderer Stelle weithin umstrittene und daher in ihrem Erlass lange unsichere DSM-RL nicht warten wollen. Man nutzte stattdessen als Grundlage die allgemeinen Ausnahmen und Beschränkungen zugunsten der Wissenschaft (Art. 5, hier insb. Abs. 3 Buchst. a) zu den 2001 in der sogenannten InfoSoc-RL 2001/29/EG ebenfalls EU-weit einheitlich geregelten Vervielfältigungsrechten von Urhebern und Leistungsschutzberechtigten (ebd., Art. 2), um soweit hierüber möglich eine erste Schranke für ein zulässiges Text und Data Mining einzuführen. Schon jetzt darf man im Rahmen von § 60d UrhG vieles rechtssicher tun.²³ Die auf der Grundlage europäischen Rechts zu erwartende Neuregelung des § 60d in Verbindung mit einem neuen § 44b UrhG-E stellt allerdings vieles klar: Insbeson-

dere wird erstmals eine Legaldefinition des Begriffs „Text und Data Mining“ aufgenommen. Zudem enthalten die beiden Regelungen Klarstellungen über Rechte und Pflichten bei der Nutzung von Werken zu Text und Data Mining sowie Schutzmechanismen für die Rechteinhaber. Die Stärke dieses künftigen Rechts ist demnach gerade die Schaffung von Rechtssicherheit für alle Beteiligten, also Rechteinhabern, Wissenschaft und Bibliotheken.

Zusammenfassend möchten wir auf zwei wesentliche Punkte hinweisen:

Der erste Punkt ist die besagte ausdrückliche Privilegierung der Kulturerbeinrichtungen einschließlich der Bibliotheken. Entscheidend ist insoweit weniger, dass das neue Recht den Bibliotheken im Bereich des Text und Data Mining mehr erlauben wird. Denn derzeit gilt ja: „Auf die gesetzliche Erlaubnis kann sich jeder berufen, der automatisierte Forschung in der oben beschriebenen Art betreiben will“²⁴ – also auch Bibliotheken. Aber die Adressierung ist eben eine völlig andere. In der Begründung des UrhWissG tauchen die Bibliotheken im Kontext des Text und Data Mining nur in ihrer klassischen Funktion als Dienstleister für Vor-, Zu- und Nacharbeiten auf.²⁵ Jetzt werden sie als Akteure angesprochen. Das ist keine Kleinigkeit. Denn „nur“ die Kulturerbeinrichtungen werden mit diesem Privileg ausgestattet anstelle vieler anderer potentiell privilegierungsfähiger Akteursgruppen wie z. B. sonstige Behörden, Journalisten, schulische Bildung, Kulturvermittlung usw., die auf viel engere rechtliche Rahmenbedingungen verwiesen werden (vgl. Art. 4 DSM-RL, § 44b UrhG-E). Es entsteht ein neuer Spielraum, eine andere Qualität von Legitimation, aber eben zugleich auch ein öffentlich sichtbarer Auftrag, sich selbst Gedanken über die inhaltliche Agenda zu machen und insofern proaktiv zu agieren im Bereich

22 Vgl. an rechtswissenschaftlicher Literatur zu UrhWissG und § 60d UrhG neben den Standardkommentaren und -Lehrbüchern zum Urheberrecht (vgl. stellv. die Kommentierungen von § 60d UrhG in Thomas Dreier/Gernot Schulze: *Urheberrechtsgesetz*, 6. Aufl., C.H. Beck: München 2018; Axel Nordemann/Jan Bernd Nordemann/Christian Czychowski (Hrsg.): *Fromm/Nordemann. Urheberrecht*, 12. Aufl., Verlag W. Kohlhammer: Stuttgart 2018; Hartwig Ahlberg/Horst-Peter Götting (Hrsg.): *BeckOK Urheberrecht*, 26. Aufl., C.H. Beck: München 2019; Artur-Axel Wandtke/Winfried Bullinger (Hrsg.): *Praxiskommentar Urheberrecht*, 5. Aufl., C.H. Beck: München 2019; Gerald Spindler/Fabian Schuster (Hrsg.): *Recht der elektronischen Medien*, 4. Aufl., C.H. Beck: München 2019) insbesondere Gerald Spindler: „Text und Data Mining – urheber- und datenschutzrechtliche Fragen“, in: GRUR 118/11 (2016), S. 1112–1120; Christian Berger: „Urheberrecht in der Wissensgesellschaft“, in: GRUR 119/10 (2017), S. 953–964; Katharina de la Durantay: „Neues Urheberrecht für Bildung und Wissenschaft – eine kritische Würdigung des Gesetzentwurfs“, in: GRUR 119/6 (2017), S. 558–567; Benjamin Raue: „Das Urheberrecht der digitalen Wissens(schafts)gesellschaft“, in: GRUR 119/1 (2017), S. 11–19; Benjamin Raue: „Text und Data Mining“, in: CR 34/10 (2017), S. 656–662; Haimo Schack: „Das neue UrhWissG – Schranken für Unterricht, Wissenschaft und Institutionen“, in: ZUM 61/11 (2017) S. 802–808; Thomas Pflüger/Oliver Hintze: „Das Urheber-Wissensgesellschafts-Gesetz aus Sicht von Hochschulen und Bibliotheken“, in: ZUM 62/3 (2018), S. 153–161; Louisa Specht: „Die neue Schrankenregelung für Text und Data Mining und ihre Bedeutung für die Wissenschaft“, in: *Ordnung der Wissenschaft* 4 (2018), S. 285–289, <http://www.ordnungderwissenschaft.de/2018-4/gesamt/37_2018_4_Specht_Schrankenregelung_fuer_Text_und_Data_Mining_odw.pdf>; Gerald Spindler: „Text und Datamining im neuen Urheberrecht und in der europäischen Diskussion“, in: ZGE 10/3 (2018), S. 273–300; Arne Umpeier: „Die Rolle der Bibliotheken in einem künftigen Urheberrechts-Wissensgesellschafts-Gesetz (UrhWissG)“, in: ZGE 10/3 (2018), S. 301–309.

23 Vgl. zur geltenden Rechtslage die Kommentierungen von § 60d UrhG in Thomas Dreier/Gernot Schulze: *Urheberrechtsgesetz*, 6. Aufl., München 2018; Axel Nordemann/Jan Bernd Nordemann/Christian Czychowski (Hrsg.): *Fromm/Nordemann. Urheberrecht*, 12. Aufl., Stuttgart 2018; Hartwig Ahlberg/Horst-Peter Götting (Hrsg.): *BeckOK Urheberrecht*, 26. Aufl., München 2019; Artur-Axel Wandtke/Winfried Bullinger (Hrsg.): *Praxiskommentar Urheberrecht*, 5. Aufl., München 2019; Gerald Spindler/Fabian Schuster (Hrsg.): *Recht der elektronischen Medien*, 4. Aufl., München 2019.

24 Begr. RegE BT-Drs. 18/12329, S. 41, <<https://dip21.bundestag.de/dip21/btd/18/123/1812329.pdf>>.

25 Vgl. Begr. RegE BT-Drs. 18/12329, S. 41, <<https://dip21.bundestag.de/dip21/btd/18/123/1812329.pdf>>: „Derjenige darf die notwendigen Handlungen auch von Dritten vornehmen lassen, z. B. von Mitarbeitern einer Bibliothek.“

des wissenschaftlichen Text und Data Mining, institutionsintern und extern bei der Suche nach Partnern in der Forschung. Das, was Arne Upmeyer über die künftige Rolle der Bibliotheken im Angesicht des UrhWissG betonte – „[d]en Bibliotheken kommt darin eine deutlich prominentere Rolle zu als bisher“²⁶ – gilt im Bereich des Text und Data Mining mit dem anstehenden neuen Recht nun also erst recht. Dies bedeutet Verantwortung, der Bibliotheken in diesem Lichte gerecht werden müssen.

Der zweite anstehende Aspekt, auf den wir hinweisen möchten, ist potentiell noch weitreichender.

Denn die Text-und-Data-Mining-fähigen Korpora, die zusammengestellt und vorbearbeitet werden aus Ursprungsmaterial, zu dem man rechtmäßigen Zugang hat, dürfen nun auch für wissenschaftliche Anschlussforschung aufbewahrt und nachgenutzt werden. Im derzeit geltenden Recht sind diese Text-und-Data-Mining-fähigen Korpora nach Projektende grundsätzlich zu löschen und dürfen allenfalls zur Überprüfung wissenschaftlicher Qualität bei bestimmten Institutionen aufgehoben und verwendet werden.

Insoweit ist das neue Recht ein Treiber, nicht nur für die Digital Humanities, sondern auch für die Bibliotheken in ihrem Kerngeschäft. Text und Data Mining hat zwar stets zum Zwecke wissenschaftlicher Forschung zu erfolgen, um in den Genuss dieses neuen Privilegs zu kommen. Aber es darf darüber hinaus durchaus zugleich gedächtnisinstitutionelle Zwecke verfolgen und fördern, d. h. solche der Erschließung, Informationsvernetzung, Auffindbarkeit, Langzeitarchivierung, Kulturvermittlung usw.

An dieser Stelle wird es positiven Druck geben. Denn die Korpora bleiben für eine etwaige wissenschaftliche Nachnutzung und gegebenenfalls Weiterentwicklung an jene Institution (oder Institutionen bei – zulässiger – Fusion von Ursprungsmaterial in einen gemeinsamen Korpus) gebunden, die den rechtmäßigen Zugang zum Ursprungsmaterial hatten. Das ergibt sich zwingend aus der Systematik des neuen Rechts, das anders als das geltende Recht nun auch ausdrücklich alles an die Voraussetzung des rechtmäßigen Zugangs binden wird (vgl. § 44b Abs. 2 S. 1 UrhG-E). D. h., dass die Digital Humanities ein erhebliches Interesse haben werden, gerade die Bibliotheken, insbesondere die mit den großen, wissenschaftlich spannenden Beständen, darin zu unterstützen und sie dazu zu befähigen, etwas aus diesem neuen Recht zu machen, so die Korpora hierfür nach Projektende eben nicht frei nachnutzbar

in die jeweilige wissenschaftliche Fachgemeinschaft weitergereicht werden können. Das eigentliche zusätzliche Potential dieses neuen Rechts, nämlich in einer ganz anderen Qualität aufeinander aufbauend Text-und-Data-Mining-basierte Forschung betreiben zu können, wird sich daher nur heben lassen, wenn die Kulturerbeeinrichtungen an dieser Stelle aktiv mitziehen, insoweit in vielen Fällen der rechtmäßige Zugang zum Ursprungsmaterial bei ihnen liegen wird. Dieses Potential des anstehenden neuen Rechts ist aber zentral für die Weiterentwicklung der Digital Humanities, um über Insellösungen und Einzelprojekthaftigkeit hinauswachsen zu können.

Was es nun bedarf, ist ein allgemeines Bekanntmachen dessen, was dieses anstehende neue Recht an Veränderung bringen wird, gepaart mit einem darauf aufbauenden breiten Dialog im Bibliothekswesen, aber auch zwischen Bibliothekswesen und Digital Humanities sowie zwischen Bibliothekswesen, Digital Humanities und Rechteinhabern darüber, wie man diese neuen Möglichkeiten produktiv zum Vorteil aller nutzen kann. Das ist keine unerhebliche Herausforderung. Es ist Neuland, das hier betreten werden wird. Aber die Chance, die hierin liegt, ist sicher größer als die zu meisternden Schwierigkeiten, schon angesichts des ständig wachsenden Bedarfs nach generischen digitalen Methoden und Tools eben nicht nur in der Wissenschaft, sondern auch im Bibliothekswesen, gerade im Bereich automatisierbarer und teilautomatisierbarer Tätigkeiten.²⁷ Das neue Recht bietet nun absehbar einen neuen Möglichkeitsraum, an dieser Stelle gemeinsam weiterzukommen. ■



Dorothea Zechmann

Leiterin Zentralbereich Verwaltung
Deutsche Nationalbibliothek
Adickesallee 1
60322 Frankfurt am Main
d.zechmann@dnb.de



Frédéric Döhl

PD Dr. phil. Ass. iur.
Deutsche Nationalbibliothek
Deutscher Platz 1
04103 Leipzig
f.doehl@dnb.de

²⁶ Arne Upmeyer: „Die Rolle der Bibliotheken in einem künftigen Urheberrechts-Wissensgesellschafts-Gesetz (UrhWissG)“, in: ZGE 10/3 (2018), S. 301-309, hier S. 301.

²⁷ Vgl. Arne Upmeyer: „Die Rolle der Bibliotheken in einem künftigen Urheberrechts-Wissensgesellschafts-Gesetz (UrhWissG)“, in: ZGE 10/3 (2018), S. 301-309, hier S. 303.