

## Text & Data Mining im wissenschaftlichen Kontext, § 60d UrhG

„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.“

(Legaldefinition des **§ 44b Abs. 1 UrhG**)

### **P** Urheberrechtliche Relevanz?

Die automatisierte Erhebung und Auswertung der Daten als solche stellt **keine urheberrechtlich relevante**, sondern eine neutrale Handlung dar.<sup>1</sup> Jedoch kommt es im Zusammenhang mit dem Text und Data Mining zu **Vervielfältigungen** des durchsuchten und gegebenenfalls urheberrechtlich geschützten Ausgangsmaterials, so dass ein **Bedürfnis für eine gesetzliche Nutzungserlaubnis** bestand.

### I. Anwendungsbereich:

§ 44b UrhG	§ 60d UrhG
Text und Data Mining <b>für jedermann</b> , wobei eine Verwendung zu <b>jedlichen Zwecken</b> erlaubt ist.	Text und Data Mining durch <b>Forscher- und Wissenschaftler:innen</b> zum Zwecke <b>wissenschaftlicher Forschung</b> .
Auch <b>kommerziell</b> ausgerichtete Akteur:innen	Berechtigt sind ausschließlich <b>nicht-kommerziell</b> agierende Akteur:innen
Enthält <b>keinen Zugangsanspruch</b> hinsichtlich des Ursprungsmaterials; der rechtmäßige Zugang wird vielmehr vorausgesetzt. <sup>2</sup> Eine Veröffentlichung des zu vervielfältigenden Werkes wird nicht vorausgesetzt.	Enthält <b>keinen Zugangsanspruch</b> hinsichtlich des Ursprungsmaterials; der rechtmäßige Zugang wird vielmehr vorausgesetzt. <sup>3</sup> Eine Veröffentlichung des zu vervielfältigenden Werkes wird nicht vorausgesetzt.
Nach Maßgabe des § 63 Abs. 1 UrhG ist <b>keine Quellenangabe</b> erforderlich, da dieser nicht auf § 44b UrhG verweist. Zum anderen wäre eine solche praktisch schwierig umsetzbar, da eine große Anzahl an Werken analysiert und in ein maschinenlesbares Format konvertiert wird.	Nach Maßgabe des § 63 Abs. 1 UrhG ist <b>keine Quellenangabe</b> erforderlich, da dieser nicht auf § 44b UrhG verweist. Zum anderen wäre eine solche praktisch schwierig umsetzbar, da eine große Anzahl an Werken analysiert und in ein maschinenlesbares Format konvertiert wird.
	§ 60d UrhG findet <b>keine Anwendung auf Computerprogramme</b> i.S.v. §§ 69a ff. UrhG, vgl. § 69d Abs. 4 UrhG.

### II. § 60d UrhG – Text und Data Mining zum Zwecke wissenschaftlicher Forschung

#### **BEACHTEN:**

Im Hochschulkontext erhält das Text und Data Mining insbesondere im Rahmen wissenschaftlicher Forschung Relevanz, so dass im Folgenden nahezu ausschließlich auf **§ 60d UrhG** eingegangen und auf § 44b UrhG lediglich zurückgegriffen wird, wenn es um allgemeine Definitionen oder Voraussetzungen geht.

<sup>1</sup> Vgl. Erwägungsgrund 9 der DSM-RL: <https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:32019L0790>; bereits zuvor BT-Drs. 18/12329, 40.

<sup>2</sup> BT-Drs. 19/27426, 96; Dreier/Schulze, 7. Aufl. 2022, UrhG § 44b Rn. 8.

<sup>3</sup> BT-Drs. 19/27426, 96.

### Die Norm im Einzelnen: § 60d UrhG

- **Abs. 1:** Enthält einen Verweis auf § 44b Abs. 1 und Abs. 2 S. 1 UrhG im Hinblick auf die Definition von „Text und Data Mining“ und die **Voraussetzungen, unter denen Vervielfältigungen** (§ 16 UrhG) zulässig sind. Dabei gilt für die Zweckbindung der „wissenschaftlichen Forschung“, dass das Text und Data Mining seinem Umfang nach im Einklang mit der Wissenschaftsfreiheit nach **Art. 13 EU-GRCh** vorgenommen werden muss.
- **Abs. 2 und 3:** Legen den Kreis der nach § 60d UrhG Berechtigten fest.  
 Im Einzelnen fallen darunter: **Forscher:innen und Forschungsorganisationen** i.S.v. § 60d Abs. 2 S. 2 UrhG, **Bibliotheken, Museen, Archive und Kulturerbe-Einrichtungen**. Dabei ist zu beachten, dass die Forscher:innen und Forschungseinrichtungen für die Inanspruchnahme der Norm **keine kommerzielle Ausrichtung** aufweisen dürfen, sämtliche Gewinne im Bereich wissenschaftlicher Forschung reinvestieren und im staatlichen Auftrag, sowie im öffentlichen Interesse tätig werden müssen.
- **Abs. 4:** Dieser sieht für einige der nach den Abs. 2 und 3 privilegierten und nicht-kommerziell tätigen Akteuer:innen die Möglichkeit der **öffentlichen Zugänglichmachung** der Vervielfältigungen nach **§ 19a UrhG** vor. Diese geht über bloße Vervielfältigungshandlungen hinaus und steht primär nur dem/der originären Rechteinhaber:in zu. Werke, die öffentlich zugänglich gemacht werden dürfen, sind solche nach **§ 2 UrhG, Datenbankwerke** und über einen Verweis in § 87i UrhG auch **Presseerzeugnisse**. Von dem Anwendungsbereich **ausgenommen** sind hingegen **Datenbanken und Computerprogramme**, was auf die Regelungsgegenstände der umzusetzenden EU Datenbank-RL und Computerprogramm-RL zurückzuführen ist.<sup>4</sup>
- **Abs. 5:** Hiernach ist die **Aufbewahrung** der angefertigten Vervielfältigungen (= Korpus) **zum Zwecke wissenschaftlicher Forschung** oder zur Überprüfung wissenschaftlicher Erkenntnisse zulässig, solange diese adäquat gegen unbefugte Nutzung geschützt werden.
- **Abs. 6:** Dieser räumt dem/der **Rechteinhaber:in** schließlich die Befugnis ein, **Maßnahmen** zu ergreifen, die einer **Gefährdung der Sicherheit und Integrität** der Netze und Datenbanken - bedingt durch die Vervielfältigen nach Abs. 1 - vorbeugen.

#### Gesetzlich erlaubt sind nach § 60d UrhG daher folgende Handlungen:

- a. Vervielfältigung einer Vielzahl von bereits zugänglichen<sup>5</sup> Werken („Ursprungsmaterial“), um daraus insbesondere durch Strukturierung und Kategorisierung eine **Sammlung von Rohdaten** (= Korpus) zu erstellen
- b. **(automatisierte) Auswertung des Korpus** zum Zwecke wissenschaftlicher Forschung
- c. **Zurverfügungstellung** des Korpus **an verbundene Forscher**, um die Überprüfbarkeit von Forschung zu gewährleisten (Peer Review)
- d. Das Ursprungsmaterial darf in der Regel **nicht weitergegeben oder online zugänglich gemacht** werden (Ausnahme in **Abs. 4** unter bestimmten Voraussetzungen und nur hinsichtlich nicht kommerziell agierenden Berechtigten)

<sup>4</sup> Dreier/Schulze, 7. Auflage 2022, UrhG § 60d Rn. 11.

<sup>5</sup> Schranke gewährt keinen Zugangsanspruch, so dass **§§ 95a und b UrhG** (technische Schutzmaßnahmen) nicht umgangen werden dürfen.

Vorgaben:

- Nur **nicht-kommerzielle** Nutzungszwecke
- Vergütungspflicht nach **§ 60h Abs. 1 UrhG** über die **§§ 54-54c UrhG**, wobei die Ansprüche ausschließlich von Verwertungsgesellschaften geltend gemacht werden können
- Nach Abschluss der Forschungsarbeit müssen alle Kopien, sowie das Ursprungsmaterial gelöscht werden. Korpus und Ursprungsmaterial dürfen jedoch Bibliotheken, Museen und Archiven zur dauerhaften Aufbewahrung übermittelt werden.<sup>6</sup>

III. Verhältnis zu ähnlichen Vorschriften:

§ 44b UrhG	§ 60d UrhG	§ 60c UrhG
Allgemeines Text und Data Mining für <b>jedermann</b> (auch) zu <b>kommerziellen Zwecken</b>	Text und Data Mining durch wissenschaftliche und <b>nicht-kommerziell</b> agierende Akteur:innen <b>zum Zwecke wissenschaftlicher Forschung</b>	Erfasst <b>Nutzungshandlungen zum Zweck wissenschaftlicher Forschung</b> an geschützten Werken <b>außerhalb des Text und Data Mining</b> , wie z.B. Vervielfältigungs- oder Verbreitungshandlungen.

<sup>6</sup> Dr. Till Kreutzer/Tom Hirche, in: Rechtsfragen zur Digitalisierung in der Lehre. Praxisleitfaden zum Recht bei E-Learning, OER und Open Content, Oktober 2017, S. 32.