

Lizenzen für Forschungsdaten

Volltextversion zum gleichnamigen Vortrag des Projektes FoDaKo¹ (Forschungsdatenmanagement in Kooperation)

Torsten Rathmann
Bergische Universität Wuppertal
Universitätsbibliothek
Gaußstr. 20
D-42119 Wuppertal
ORCID: 0000-0001-5880-1546

Zusammenfassung

Dieser Text erläutert, warum es wichtig ist, öffentlich zugängliche Forschungsdaten zu lizenzieren und gibt eine Übersicht über die kostenlos verfügbaren Lizenzen von Creative Commons. Dabei wird auch darauf eingegangen, welche Schwierigkeiten der Ausschluss kommerzieller Nutzung und der Veröffentlichung abgewandelter Daten mit sich bringt. Ein kurzer Abschnitt befasst sich mit den Themen, die allgemein Bestandteil einer Lizenz sein sollten, ein weiterer damit, welche Lizenzen mitgegeben werden können, wenn Forschungsdaten bei RADAR oder EUDAT gespeichert werden.

Einführung

In den letzten Jahren gab es in der öffentlichen Meinung eine Tendenz hin zu mehr Offenheit für Daten, die mit öffentlichen Geldern gesammelt werden, und für wissenschaftliche Daten im Allgemeinen. Gemeinsame rechtliche Standards für das Kopieren, die Nutzung und die Veränderung intellektueller Werke gewinnen daher immer mehr an Bedeutung, zumal viele Geldgeber ihren Einfluss auf die Datenzugriffspolitik von Projekten nutzen, um die Projektpartner zu drängen, ihre Produkte für die Nachnutzung zu öffnen.

Es gibt mehrere gute Gründe, warum sich Dateneigentümer mit den Möglichkeiten der Lizenzvergabe befassen sollten. Dateneigentümer können sich durch Vergabe einer Lizenz für ihre Forschungsdaten vor Schadenersatzansprüchen schützen und unerwünschte Arten der Nutzung verhindern. Darüber hinaus wird Datennutzern durch die Verwendung bekannter Lizenzen die Nachnutzung sogar erleichtert, denn ohne ausdrückliche Lizenzbedingungen müssen nationale Rechte an geistigem Eigentum befolgt werden und diese Rechte hängen vom jeweiligen Land ab und sind oft komplex. Last but not least erlauben viele Archive nicht alle möglichen Lizenztexte, sondern lassen Dateneigentümer einen Text aus einer Reihe von unterstützten Lizenzen auswählen.

¹ <https://www.fodako.de/>

Die CC-Lizenzen

Ein bekanntes System von Lizenzen wurde von der [Creative Commons Corporation \(CC\)](#) eingeführt. CC-Lizenzen sind weit verbreitet, kostenlos und können zum Schutz aller Arten von Inhalten verwendet werden, aber nur eingeschränkt zum Schutz der Datenbankstruktur und Sammlung² und gar nicht von Software. CC-Lizenzen sind einfach zu verwenden, da jede Lizenz aus einer rechtlich verbindlichen Anordnung, einer Zusammenfassung, einem Symbol, das als Grafikdatei heruntergeladen und in ein Dokument eingefügt werden kann, und einem maschinenlesbaren Symbol besteht.

CC hat mehrere Lizenztexte für verschiedene Zwecke veröffentlicht. [CC0](#) ist der umfassendste Verzicht auf Rechte. Der Dateneigentümer gibt alle Urheberrechte und verwandten Schutzrechte im größtmöglichen, gesetzlich zulässigen Umfang auf und gibt sie ab. Mit CC0 können Daten als *Public Domain* deklariert werden. Nachnutzer dürfen die Daten kopieren, ändern, verteilen, anzeigen und sogar für kommerzielle Zwecke verwenden, ohne um Erlaubnis bitten zu müssen³.

Mit der CC BY-Lizenzfamilie sind mehr Pflichten verbunden. Wird Material weitergegeben, das durch eine solche Lizenz geschützt ist, muss Folgendes beigefügt werden:

- der Name des Erstellers und anderer, die für eine Namensnennung vorgesehen sind,
- ein Copyright-Vermerk,
- einen Hinweis auf den Lizenztext,
- einen Hinweis auf den Haftungsausschluss und
- einen Link zum lizenzierten Originalmaterial.⁴

Eine Übersicht über die CC BY-Lizenzfamilie finden Sie in Abbildung 1. [CC BY](#), die Creative-Commons-Lizenz mit Namensnennungspflicht und dem ansonsten weitgehensten Rechteeverzicht, ist natürlich die Grundlage. Die anderen Lizenzen enthalten zusätzliche Regelungen, um die kommerzielle Nutzung zu verbieten oder wie mit abgeleiteten Arbeiten, z.B. Übersetzungen oder Änderungen, umzugehen ist. [CC BY-ND](#) erlaubt keine Veröffentlichung von abgewandeltem Material, [CC BY-SA](#) hingegen erlaubt sie, aber nur unter derselben oder einer kompatiblen Lizenz (ShareAlike). Welche Lizenzen kompatibel sind, entscheidet Creative Commons. CC BY selbst erlaubt die Veröffentlichung von abgewandeltem Material auch unter einer anderen Lizenz.

Im rosafarbenen äußeren Ring von Abbildung 1 befinden sich die drei Lizenzen, die keine kommerzielle Verwendung erlauben. CC verwendet die folgende Definition von NonCommercial⁵.

² Licence and Legal FAQ/Why CC BY-SA is Unsuitable, https://wiki.osmfoundation.org/wiki/Licence_and_Legal_FAQ/Why_CC_BY-SA_is_Unsuitable, aufgerufen am 10.10.2018

³ <https://creativecommons.org/publicdomain/zero/1.0/legalcode>, aufgerufen am 10.10.2018

⁴ <https://creativecommons.org/licenses/by/4.0/legalcode>, aufgerufen am 10.10.2018

⁵ Beispielsweise <https://creativecommons.org/licenses/by-nc/4.0/legalcode.de>, aufgerufen am 12.10.2018

Nicht kommerziell meint nicht vorrangig auf einen geschäftlichen Vorteil oder eine geldwerte Vergütung gerichtet. Der Austausch von lizenziertem Material gegen anderes unter Urheberrecht oder ähnlichen Rechten geschütztes Material durch digitales File-Sharing oder ähnliche Mittel ist nicht kommerziell im Sinne der vorliegenden Public License, sofern in Verbindung damit keine geldwerte Vergütung erfolgt.

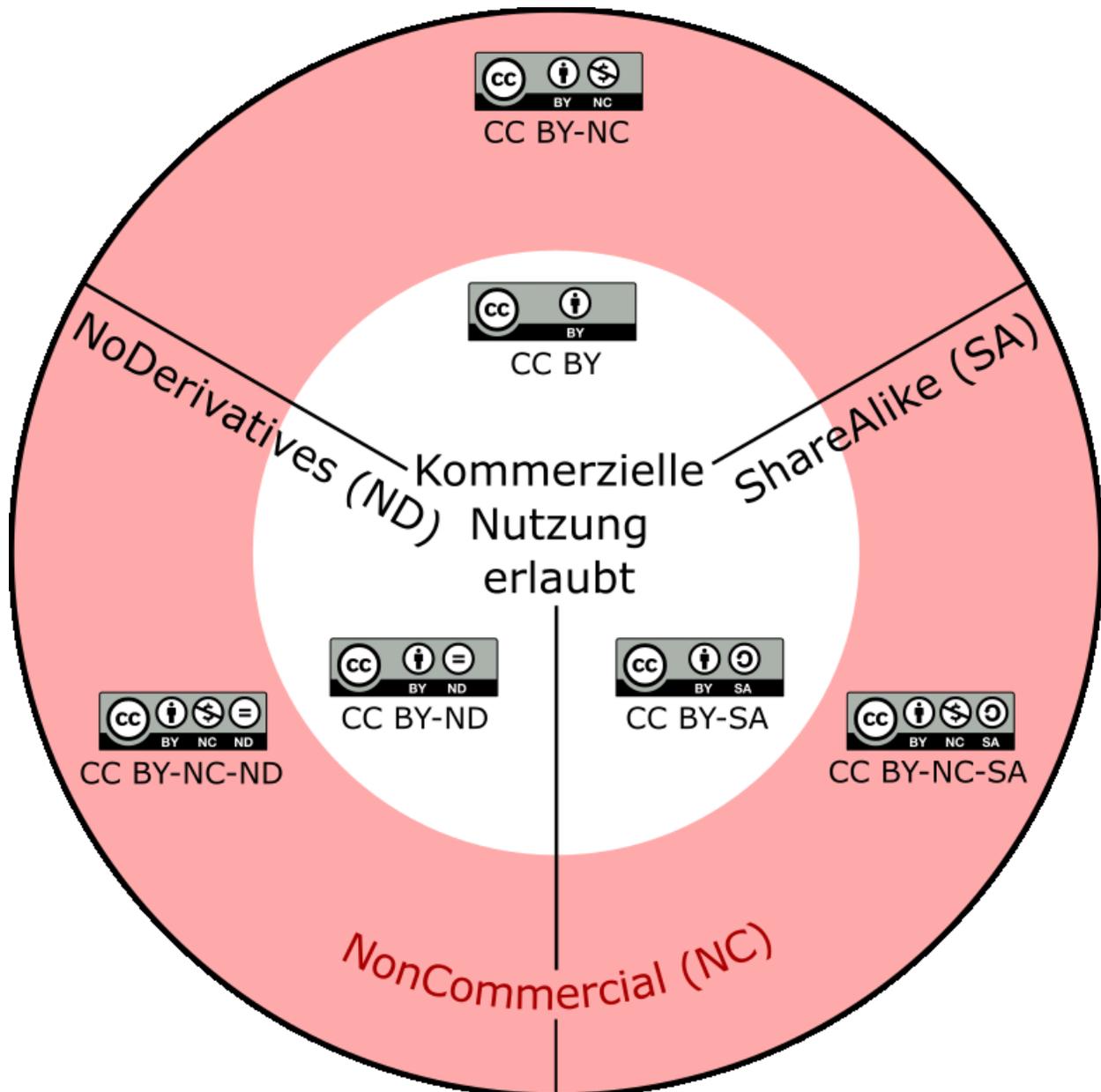


Abbildung 1: Übersicht über die die Creative-Commons-Lizenzen mit Namensnennungspflicht

[CC BY-NC](#) schließt die kommerzielle Nutzung aus, erlaubt aber die Weitergabe von abgewandeltem Material. [CC BY-NC-ND](#) erlaubt weder die kommerzielle Nutzung noch die Weiter-

gabe von abgewandeltem Material. [CC BY-NC-SA](#) verbietet kommerzielle Nutzung, erlaubt aber die Weitergabe von abgewandeltem Material unter derselben Lizenz.

Problematisch: Ausschluss kommerzieller Nutzung

Auf der einen Seite ist es verständlich, dass viele Forscher irritiert sind, wenn ihre Daten, die in öffentlich finanzierter Forschung produziert wurden, von anderen verkauft werden, z.B. von einem Unternehmen, das damit Geld verdient. Auf der anderen Seite kann das Verbot der monetären Kompensation der Mitarbeitergehälter aber verhindern, dass Nichtregierungsorganisationen (NGOs) die Daten nachnutzen. Nur selten können sich größere Organisationen allein auf ehrenamtliche Arbeit stützen. Ein Teil der NGO-Arbeit wird üblicherweise von Mitarbeitern mit normalen Arbeitsverträgen geleistet. Die geldwerte Vergütung (monetäre Kompensation) der bezahlten Arbeitsstunden durch Einnahme von Gebühren, z.B. für den Vertrieb, ist aber unter CC BY-NC verboten.

Problematisch: Ausschluss der Weitergabe abgewandelter Materials

CC BY-ND und CC BY-NC-ND erlauben keine Weitergabe von abgewandeltem Material. Der entscheidende Abschnitt der CC BY-ND-Lizenz lautet wie folgt⁶:

Unter den Bedingungen der vorliegenden Public License gewährt der Lizenzgeber Ihnen eine weltweite, vergütungsfreie, nicht unterlizenzierbare, nicht-ausschließliche, unwiderrufliche Lizenz zur Ausübung der lizenzierten Rechte am lizenzierten Material, um:

1. das lizenzierte Material ganz oder in Teilen zu vervielfältigen und weiterzugeben; und
2. abgewandeltes Material zu erstellen und zu vervielfältigen, es aber nicht weiterzugeben.

Creative Commons definiert *abgewandeltes Material* wie folgt⁶:

Abgewandeltes Material bezeichnet Material, welches durch Urheberrechte oder ähnliche Rechte geschützt ist und vom lizenzierten Material abgeleitet ist oder darauf aufbaut und in welchem das lizenzierte Material übersetzt, verändert, umarrangiert, umgestaltet oder anderweitig modifiziert in einer Weise enthalten ist, die aufgrund des Urheberrechts oder ähnlicher Rechte des Lizenzgebers eine Zustimmung erfordert.

Der Ausschluss von abgewandeltem Material ist eine ernsthafte Einschränkung.

1. Das Recht, aus einem Werk Ableitungen zu ziehen, ist eines der zentralen Rechte der kulturellen Freiheit⁷.

⁶ <https://creativecommons.org/licenses/by-nd/4.0/legalcode.de>, aufgerufen am 12.10.2018

⁷ <http://freedomdefined.org/Definition>, aufgerufen am 10.10.2018

2. Abgewandeltes Material ist häufig ein wichtiger wissenschaftlicher Beitrag, z.B. in der Klimaforschung: Der Output von globalen Simulationen wird als Input für regionale Klimamodelle verwendet. Regionale Klimamodelle haben in der Regel eine bessere Auflösung als globale Modelle und können Ergebnisse liefern, die für lokale Gemeinschaften wichtig sind. Regionale Klimamodelle unterscheiden sich von globalen Modellen und die Entwicklung und Anwendung regionaler Modelle ist eine eigenständige, vollwertige wissenschaftliche Arbeit.
3. Da Original und Abwandlung wesentlich voneinander abweichen müssen, ist nicht jede kleine Änderung eine Abwandlung. Dennoch ist die Festlegung einer Grenze zwischen einer kleinen Änderung und einer Abwandlung schwierig. Diese Unsicherheit kann die Weitergabe geringfügiger Änderungen wie Datenkorrekturen behindern.

Forscher sollten überlegen, ob Abwandlungen wirklich ausgeschlossen werden sollen.

Unabhängig davon, ob Abwandlungen ausgeschlossen sind oder nicht, werden Forscher, die Daten erneut veröffentlichen wollen, die durch eine der CC BY-Lizenzen geschützt sind, aufgefordert anzugeben, ob sie *das lizenzierte Material verändert haben, und alle vorherigen Änderungsangaben beibehalten* haben. Diese Regel soll sicherstellen, dass die geänderten Daten klar vom Original unterscheidbar sind.

Problematisch: ShareAlike-Lizenzen

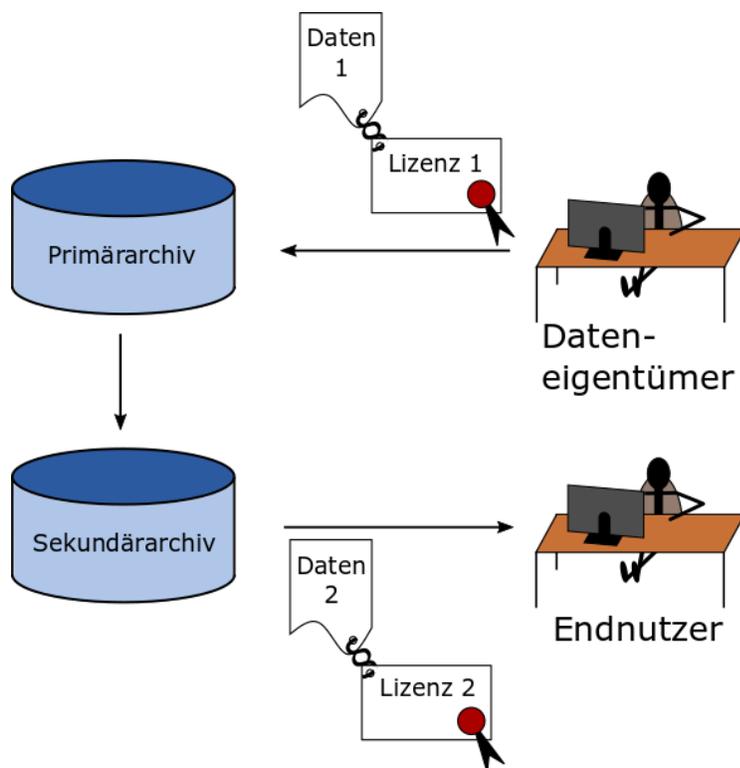


Abbildung 2: Lizenzen im Falle einer Wiederveröffentlichung von Forschungsdaten. ND (NoDerivatives) heißt Daten 1 \approx Daten 2 (nur sehr kleine Änderungen erlaubt). SA (ShareAlike) heißt Lizenz 1 \approx Lizenz 2 (Lizenzen müssen gleich oder kompatibel sein)

CC ShareAlike-Lizenzen ermöglichen die Weitergabe von Abwandlungen unter derselben oder einer kompatiblen Lizenz. In einem Beispiel, das in Abbildung 2 gezeigt ist, nimmt ein Archiv (sekundäres Archiv) Daten von einem anderen Archiv (primäres Archiv) und bietet sie an, vielleicht zusammen mit zusätzlichen Diensten auf den Daten. Falls die Originaldaten durch eine CC BY-NC-SA-Lizenz geschützt sind, muss das Sekundärarchiv den Daten dieselbe Lizenz zuweisen, unabhängig davon, ob die Daten geändert wurden oder nicht und welche zusätzlichen Dienste mit den Daten angeboten werden. In diesem Fall kann das Sekundärarchiv keine finanzielle Entschädigung von kommerziellen Nutzern erhalten, z.B. für teure zusätzliche Qualitätssicherung oder Dienstleistungen. Da das dazu führen kann, dass das Sekundärarchiv die Daten und Dienste überhaupt nicht anbietet, sollten Dateneigentümer überlegen, ob dies so beabsichtigt ist.

Warum ist eine Lizenz für Nachnutzer so wichtig?

Dateneigentümer sollten ihren veröffentlichten Daten eine Lizenz zuweisen, da andernfalls nationale Immaterialgüterrechte gelten würden, wenn die Lizenz fehlt. Nationale Rechte an geistigem Eigentum haben zwei Nachteile: Sie unterscheiden sich von Land zu Land, und nationales Recht kann sehr ländlich sein, in einer fremden Sprache geschrieben und für Nicht-Juristen schwer zu interpretieren sein. An den speziellen Fall angepasste kurze Nutzungsbedingungen sind normalerweise die bessere Lösung.

Was sollte unbedingt in die Lizenz hinein?

Die Lizenz sollte enthalten:

- Haftungsausschluss und Haftungsbeschränkung
- Namensnennungspflicht
- Regeln für Datenbanken und Sammlungen

Alle Lizenzen der CC BY-Familie erfüllen diese Anforderungen. Der letzte Punkt ist wichtig, wenn der Dateneigentümer über *sui generis*-Datenbankrechte verfügt. In diesem Fall können die Rechte für die Daten selbst und die Rechte für die Datenbank und die Sammlung in der Datenbank getrennt werden. Um eine Nutzung der Daten zu ermöglichen, sollte der Dateneigentümer dem Benutzer auch erlauben, Inhalte der Datenbank zu extrahieren, wiederzuverwenden und zu teilen.

Lizenzen und externe Speicherung

Dateneigentümer, die ihre Forschungsdaten in einem externen Rechenzentrum speichern wollen, müssen prüfen, ob die Lizenz, die die Daten schützen soll, und die Nutzungsbedingungen des Archivs kompatibel sind. Beispielsweise sind Forschungsdaten mit eingeschränktem Zugang in einigen Archiven nicht willkommen. Für viele Forscher könnten die Dienste von RADAR oder EUDAT nützlich sein, da RADAR und EUDAT für alle Fächer offen sind.

EUDAT

Das europäische Projekt [EUDAT](#) bietet eine ganze Reihe von Diensten an. Interessant für Dateneigentümer und Datennutzer sind die Dienste in der folgenden Liste.

- **B2FIND:** Für diesen großen Katalog werden nur die Metadaten von den Datenanbietern übernommen. Alle Metadaten sind vollständig offen. Dies wird zum Zeitpunkt des ersten Kontakts zwischen dem Rechenzentrum und dem Metadatenanbieter vereinbart. Andere Metadaten sind nicht erwünscht und für eine Metadaten-Suchmaschine wie B2FIND auch nicht sinnvoll. Im Gegensatz zu den Metadaten müssen die Daten nicht offen sein. Der B2FIND-Katalog enthält nur Metadaten und Links zu den Datenanbietern. Von daher besteht keine Notwendigkeit, mit den geistigen Eigentumsrechten an den Daten selbst umzugehen.
- **B2SHARE:** Dieser Dienst bietet einen Daten-Upload für offene Forschungsdaten mit geringem Volumen (Long-Tail-Daten, d.h. meist schon verarbeitete und damit aggregierte Daten) an. Die Kunden müssen ihre Daten selbst hochladen. Während des Hochladens können sie aus einer Liste verschiedener standardisierter offener Lizenzen auswählen oder einen eigenen Lizenztext hochladen. Der Lizenztext selbst muss nicht unbedingt frei nachnutzbar sein.
- **B2SAVE:** Die Speicherung von Daten mit hohem und mittlerem Volumen, die von diesem EUDAT-Dienst angeboten wird, erfordert eine gewisse Sorgfalt bezüglich der Rechte an geistigem Eigentum. EUDAT hat selbst nicht den notwendigen Speicher für solche Datenvolumina und arbeitet stattdessen mit Rechenzentren zusammen, die den Speicher zur Verfügung stellen. Interessierte Dateneigentümer sollten Kontakt aufnehmen für den Fall, dass die Daten nicht vollständig offen sind. Nicht jedes Rechenzentrum akzeptiert jede Lizenz, denn der Umgang mit einer Vielzahl von Lizenzen ist eine Last für die Rechenzentren.⁸

RADAR (Research Data Archive, FIZ Karlsruhe)

[RADAR](#) bietet die Speicherung von Forschungsdaten als kostenpflichtige Dienstleistung für alle Fächer an. Da alle Datenformate akzeptiert werden, auch proprietäre und nicht gut dokumentierte Formate, wird nur die Bitstream-Erhaltung angeboten, keine Datenpflege. RADAR akzeptiert Daten, die

- nicht veröffentlicht werden sollen
- veröffentlicht werden sollen unter
 - CC0
 - einer der Lizenzen der CC BY-Familie
 - © Copyright, d.h. unter deutschem Urheberrecht

⁸ https://portal.enes.org/ISENES2/documents/deliverables/is-enes2_d5-3_report-on-basic-data-access-protocols-and-data-quality-control/view, aufgerufen am 10.10.2018

Weiterführende Links

<https://www.forschungsdaten.info/themen/rechte-und-pflichten/lizenzvergabe/>

<https://creativecommons.org/>



Dieses Werk ist lizenziert unter
einer [Creative Commons Namens-
nennung 4.0 International Lizenz](https://creativecommons.org/licenses/by/4.0/).

Dieser Aufsatz ist im Wesentlichen eine deutsche Übersetzung der FoDaKo-Information
Licenses for Research Data, <https://doi.org/10.5281/zenodo.1463156>