

DATENLIZENZIERUNG IN DER CLOUD – STATUS QUO UND IMPLIKATIONEN

Tassilo Pellegrini / Ivan Ermilov

Professor (FH), Fachhochschule St. Pölten, Department Medienwirtschaft
Matthias Corvinus Straße 15, 3100 St. Pölten, AT
Tassilo.pellegrini@fhstp.ac.at; <http://www.fhstp.ac.at>

Doktoratsstudent, Universität Leipzig, Institut für Informatik, Deutschland
Augustusplatz 10, 04109 Leipzig, DE
iermilov@informatik.uni-leipzig.de; <http://aksw.org/About.html>

Schlagworte: *Datenlizenzierung, Daten-Portale, Cloud Computing, Derivative Werke, Intellectual Property Rights, IRIS 2016*

Abstract: *Mit der zunehmenden technischen Verfügbarkeit hochwertig strukturierter Daten über Cloud-Anwendungen und Portale stellt sich die Frage, wie diese Ressourcen lizenzrechtlich in die digitale Wertschöpfung integriert werden können. Eine Herausforderung für das Rechtemanagement stellen hierbei derivative Werke dar. Einem grundsätzlichen Problembewusstsein steht eine dünne Befundlage zum Status quo und den Folgeproblemen der Lizenzierungspraxis in Daten-Clouds gegenüber. Der Beitrag versucht diese Lücke teilweise zu schließen. Auf Basis der Befunde werden Herausforderungen und Perspektiven für die Datenbewirtschaftung unter netzökonomischen Bedingungen illustriert.*

1. Einleitung

Mit der steigenden Interoperabilität von IT-Systemen und -Plattformen sowie der damit verbundenen Portabilität von Daten rücken neben Content auch technische Artefakte wie Daten und korrespondierende Services in den Fokus des kommerziellen Interesses. Dies zeigt sich einerseits in der wachsenden Bedeutung von Daten in der digitalen Wertschöpfung – speziell im Kontext der stark auf Datenanalytik, Automatismen und Co-Produktion basierenden service-orientierten Logik des eCommerce –, andererseits in der fortwährenden Entwicklung neuer Standards und Methoden zur Bewirtschaftung von Daten unter netzökonomischen Bedingungen [SHY 2001, S. 10]. Beispiele dafür sind etwa die Semantic Web Initiative des World Wide Web Konsortiums¹, das Projekt «WikiData» der Wikimedia Foundation² oder die Metadateninitiative «schema.org»³ der Suchmaschinenbetreiber Google, Bing und Yahoo. Begriffe wie «Big Data» als konzeptioneller Überbau und «Linked Data XE» als technologisch-methodische Basis sind Ausdruck einer zunehmenden Professionalisierung der Datenbewirtschaftung und damit einhergehender Kommerzialisierung. Befeuert wird diese Entwicklungen durch die partielle bis totale Offenlegung von Daten im Zuge von Open Access bzw. Open Government Initiativen oder Open Innovation Strategien, sowie den zunehmenden Einsatz von Dual Licensing Strategien als Bestandteil einer geschäftsimmanenten Logik in der Bewirtschaftung vernetzter Daten.

Mit der Entwicklung sogenannter Daten-Clouds [JENTZSCH 2014, S. 209], d.h. der Verfügbarmachung großer Datenmengen und dazugehöriger Dienste mittels technischer Schnittstellen über das Internet, können Daten

¹ Für einen Überblick über Semantic Web Standards siehe <http://www.w3.org/standards/semanticweb/> (aufgerufen am 16. August 2015).

² Siehe https://www.wikidata.org/wiki/Wikidata:Main_Page (aufgerufen am 16. August 2015).

³ Siehe <https://schema.org/> (aufgerufen am 16. August 2015).

mit relativ geringem Aufwand aus unterschiedlichen Quellen bezogen, in bestehende Prozesse integriert und zu neuen Produkten kombiniert werden. Damit gewinnen auch Fragen der rechtssicheren Verwendung im Sinne einer «legal compatibility» [HOFFMANN ET AL. 2015] von Daten an Bedeutung, was insbesondere dort relevant ist, wo Daten unterschiedlicher Rechteinhaber zu derivativen Werken kompiliert und vermarktet werden.

Die Nutzung von Daten wird basierend auf dem Urheber- und Datenbankrecht mittels Lizenzen geregelt. Diese definieren die Nutzungsbedingungen, unter denen Daten für kommerzielle und nicht-kommerzielle Zwecke einer Bewirtschaftung zugeführt werden können. Neben dem Urheber- und Datenbankrecht, die dem Rechteinhaber ein temporäres Monopol auf die Verwertung seiner geistigen Schöpfung einräumen, kommen zunehmend auch offene Lizenzen zum Einsatz, die mit einem totalen oder partiellen Verzicht auf (symbolische oder monetäre) Kompensation des Rechteinhabers einhergehen.⁴ Nichtsdestotrotz können Lizenz- und Nutzungsbestimmungen einzelner Datensätze (teilweise stark) voneinander abweichen und zu Kompatibilitätskonflikten führen, was deren (kommerzielle) Wiederverwertbarkeit einschränkt und gegebenenfalls Zusatzkosten bis hin zu juristische Folgeprobleme verursachen kann.

Die Generierung derivativer Werke kann u.U. mit erheblichem Rechtklärungsaufwand und damit verbundenen Transaktionskosten einhergehen. Diese lassen sich in der Phase der Rechtklärung folgendermaßen einteilen:

1. Anbahnungskosten: Recherche und Vorauswahl der geeigneten Datenquellen
2. Informationsbeschaffungskosten: Recherche ausreichender Lizenzinformation für den konkreten Anwendungsfall
3. Vereinbarungskosten: Prüfen von Lizenzkonflikten und ggfs. Aushandlung individueller Nutzungsbedingungen

In Anlehnung an die Definition der Vereinten Nationen sei im Folgenden unter «Datensatz» verstanden, jegliche systematische, schützenswerte Kompilation von Daten, die in elektronischer Form zugänglich und als eigenständiges Werk identifizierbar ist.⁵

Unter «derivativem Werk» soll in Anlehnung an die Definition der World Intellectual Property Organization verstanden werden, ein Werk, das aus der Übersetzung, Adaption, Kompilation (oder vergleichbaren Verfahren) eines oder mehrerer vorhergehender Werke generiert wurde.⁶

Bis dato gibt es kaum empirische Untersuchungen zur effektiven Situation der Verfügbarkeit und Güte geschäftsrelevanter Lizenzinformationen in bestehenden Daten-Clouds. Diese Forschungslücke soll mittels einer Querschnittsbetrachtung der aktuellen Lizenzierungspraxis auf vier ausgewählten Datenportalen teilweise geschlossen werden. Die vorliegenden Befunde sollen dabei helfen, Folgeprobleme zu identifizieren und Entwicklungsperspektiven des digitalen Rechtemanagements unter Bedingungen cloud-basierter Datenbewirtschaftung zu diskutieren.

⁴ Exemplarisch genannt seien etwa Creative Commons, Open Data Commons und das gesamte Lizenzspektrum unter dem Label Free- und Open Source Software. Einen Überblick über verfügbare Open Source und Open Commons Lizenzen findet sich bei <https://tldrlegal.com/> (aufgerufen am 16. August 2015).

⁵ Siehe <https://data.un.org/Glossary.aspx?q=data> (aufgerufen am 16. August 2015).

⁶ Siehe <http://www.wipo.int/tk/en/resources/glossary.html#2> (aufgerufen am 16. August 2015).

2. Datenlizenzierung und Lizenzierungspraxis – Forschungsstand

In den vergangenen Jahren hat sich ein generelles Bewusstsein zur Problematik des Rechte- und Lizenzmanagements im Kontext derivativer Werke [GROVES 1997] und der Schutzwürdigkeit von Daten [SONNTAG 2006] entwickelt. Eine Intensivierung erhielt die Diskussion im Umfeld der Open Source Bewegung [VÄLIMÄKI 2005], bei der im Zuge der kommerziellen Adaption von quelloffener Software Folgeprobleme der offenen Lizenzierung [STALLMAN ET AL. 2006] virulent wurden. Aktuellere Arbeiten zum Thema erweitern den Betrachtungsfokus von der Software- auf die Content-Produktion [GUIBAULT & ANGELOPOULOS 2011], lenken den Fokus auf Fragen des Wertschöpfungsbeitrages der Datenbewirtschaftung [PELLEGRINI 2012] und kritisieren rechtliche Lücken in der Lizenzierungsgebarung bei der Erstellung derivativer Werke [MAZZIOTTI 2008]. Vor dem Hintergrund der zunehmenden Automatisierung der Content-Verarbeitung und der wachsenden Bedeutung multimedialer Zusatzangebote im Kontext der Mehrfachverwertung von Medienangeboten erfährt die Diskussion des strategischen [SCHAAAL 2010] und technischen [BACKHAUS 2008] Rechtemanagements eine Verbreiterung.

Explizite Fragen zum Problem der Datenlizenzierung und Vorschläge für die Ausgestaltung von Dual Licensing Policies [FROSTERUS ET AL. 2011, HYLAND 2011, PELLEGRINI 2014] tauchen im Zuge der zunehmenden Verfügbarkeit offener Daten als Resultat von Open Government Data [ARCHER ET AL. 2013] und Open Access [HYVONEN 2012] Initiativen auf. Komplementär dazu findet sich eine wachsende Anzahl von Beiträgen, die sich Fragen der maschinenlesbaren Repräsentation und automatischen Prozessierung von Lizenzinformation mittels Reasoning- und Inferenzverfahren widmet [PUCELLA & WEISSMAN 2002, GARCIA & GIL 2009, VILLATA & GANDON 2012, ROTOLO ET AL. 2013].

Generell liegt diesen Ausführungen eine schlechte empirische Datenlage zugrunde, insbesondere im Kontext cloud-basierter Lizenzierungspraxis. Eine Erhebung der Ontology Engineering Group der Universität Madrid im Jahr 2013 zur Verfügbarkeit von Lizenzinformationen in der sog. «Linked Open Data Cloud» ergab, dass etwa 40% aller Datensätze keine Angaben zu den Nutzungsbedingungen enthielten bzw. in etwa 30% der Fälle die verfügbare Lizenzinformation nicht das volle Schutzspektrum vor allem im Zusammenspiel aus Urheber- und Datenbankrecht bediente.⁷ Zu vergleichbaren Befunden kommen auch JAIN ET AL. [2013]. In einer aktuelleren Untersuchung der Linked Open Data Cloud durch JENTZSCH [2014] wurden neben dem thematischen Spektrum der verfügbaren Daten sowohl deren technische Güte als auch die verfügbaren Lizenzinformationen in maschinen-verarbeitbarer Form untersucht. Laut ihren Ergebnissen weisen gar nur 2,7% der verfügbaren Datensätze eine maschinenlesbare Referenz zu Lizenzinformationen aus. Eine cloud-vergleichende Analyse zur Lizenzierungspraxis ist Bestandteil der folgenden Ausführungen.

3. Befunde und Folgeprobleme der Lizenzierungspraxis

3.1. Fragestellungen und Methode

Um die Kenntnislage zur effektiven Lizenzierungspraxis und daraus resultierender Probleme zu verbessern, wurden vier Datenportale, die als Access Points zu unterschiedlichen Daten-Clouds dienen, einer Analyse unterzogen. Im Kern der Untersuchung standen folgende Fragen:

1. Sind die Datensätze mit Lizenzinformation ausgestattet?
2. Welche Lizenzen kommen zum Einsatz?
3. Existieren Kompatibilitätskonflikte innerhalb und zwischen den Daten-Clouds?

⁷ Siehe auch <http://oeg-dev.dia.fi.upm.es/licensius/static/observatory/##/step-1> (aufgerufen am 16. August 2015).

4. Wie hoch ist der Anteil maschinen-verarbeitbarer Lizenzinformation?

Untersucht wurden drei Open Government Data Portale sowie ein Open Science Portal. Bei den Open Data Government Portalen handelt es sich um data.gov (USA)⁸, OpenCanada (Kanada)⁹ und publicdata.eu (EU)¹⁰. Die Auswahl erlaubt den Vergleich regionaler Unterschiede in der Lizenzgebung. Zusätzlich wurde datahub.io analysiert, ein offenes, überregionales Datenrepositorium der Open Knowledge Foundation¹¹, welches vorwiegend für die Veröffentlichung wissenschaftlicher Daten Verwendung findet.

Die Analyse erfolgte im April 2015 auf Basis einer automatischen Extraktion der Lizenzinformation von den Landing Pages der zugänglichen Datensätze mittels eines eigens programmierten Web-Crawlers.¹² In Summe konnten mit dieser Methode 441.315 Datensätze auf die Verfügbarkeit von Lizenzinformation untersucht werden und Aussagen über deren Güte und Eignung für eine maschinelle Verarbeitbarkeit getroffen werden.

		Datagov	OpenCanada	PublicData	Datahub
1	Datensätze	132.206	244.257	55.481	9.371
2	Lizenztypen	10	3	50	33
3	Creative Commons	0,4%	0,0%	35,3%	17,1%
4	Open Data Commons	0,0%	0,0%	0,5%	4,8%
5	Individuelle Lizenz	0,0%	100%	39,9%	19,0%
6	Keine Lizenz	99,6%	0,0%	24,3%	59,1%
7	Deref. Link	0,4%	100%	43,2%	23,1%
8	Mach. Read.	0,0%	0,0%	2,6%	2,2%

Tabelle: Lizenzen auf ausgewählten Datenportalen (Stand April 2015)

Erhoben wurden die Anzahl der zugänglichen Datensätze (1), die Anzahl (2) und Art (3, 4, 5) der unterschiedlichen Lizenztypen, der Anteil von Datensätzen ohne Lizenzinformation (6), der Anteil de-referenzierbarer Lizenzinformation (7) und der Anteil von Datensätzen mit maschinenlesbarer Lizenzinformation (8). Die Tabelle veranschaulicht die Analyseergebnisse im Überblick. Im Folgenden sollen einige Befunde herausgegriffen und Problemlagen identifiziert werden.

3.2. Befund 1: Unzureichende Informationslage

Die Analyse offenbart große regionale Unterschiede in der Verfügbarmachung von Lizenzinformation.

Auf data.gov weisen 99,6% der Datensätze keine Lizenzinformationen aus. Dies ist zu einem Gutteil auf den Umstand zurückzuführen, dass Daten, die von Bundesbehörden erhoben werden, unter der US Open Data Policy¹³ lizenziert werden müssen. Diese Information ist zwar in den Nutzungsbedingungen des Portals hinterlegt, jedoch auf Ebene der Datensätze nicht expliziert. Entsprechend konnte sie auch nicht automatisch extrahiert,

⁸ Siehe <http://www.data.gov/> (aufgerufen am 16. August 2015).

⁹ Siehe <http://open.canada.ca/en> (aufgerufen am 16. August 2015).

¹⁰ Siehe <http://www.publicdata.eu/> (aufgerufen am 16. August 2015).

¹¹ Siehe <https://okfn.org/> (aufgerufen am 16. August 2015).

¹² Der Crawler wurde als GitHub-Projekt veröffentlicht unter <https://github.com/AKSW/ckan-aggregator-py> (aufgerufen am 16. August 2015).

¹³ Siehe <http://www.data.gov/data-policy> (aufgerufen am 16. August 2015).

sondern muss manuell recherchiert werden. Kompliziert wird die Rechtelage zusätzlich dadurch, dass knapp 25% der verfügbaren Datensätze von Behörden anderer Verwaltungsebenen oder von staatsnahen Organisationen stammen, die nicht an die Open Data Policy gebunden sind. Nur 0,4% der Datensätze weisen explizit eine Lizenz aus und bieten einen Link zu den Vertragsbestimmungen.

Auf datahub.io verfügen 59,1% aller Datensätze über keine Lizenzinformation und nur 23,1% bieten einen referenzierbaren Link zu den Lizenzbestimmungen an. Auf publicdata.eu weisen 24,3% der Datensätze keine Lizenzinformation aus, jedoch verweisen 43,2% der Datensätzen auf einen Link zu den Lizenzbestimmungen.

Im Kontrast dazu steht das kanadische Open Government Data Portal, auf dem 100% der verfügbaren Datensätze über eine Lizenzinformation verfügen, die auch maschinell verarbeitet werden kann. Allerdings ist anzumerken, dass die Zuteilung einer Lizenz Bestandteil des Workflows ist und auf OpenCanada nur Regierungsdaten veröffentlicht werden.

Die Befunde für datahub.io und publicdata.eu können sich unterschiedlich erklären lassen. Zum einen können die nötigen Lizenzinformationen schlichtweg fehlen, weil der Datenanbieter keine Erfahrung mit bzw. keine Kenntnis von der Lizenzierungspraxis hat oder aber die nötige Information nicht adäquat im Datensatz hinterlegt hat. So zeigt sich, dass die nötigen Lizenzinformationen an unterschiedlichsten Stellen und in unterschiedlichsten Formaten – von einem Statement auf der Webseite, über einen eingebetteten XML-Kommentar, bis zu alleinstehenden .txt-Files – verfügbar gemacht werden. Zum anderen können abweichende nationale Regelungen unterschiedliche Lizenzierungskonventionen bedingen. Während nämlich in den USA Datensätze ohne ausgewiesene Lizenzinformation generell als gemeinfrei gelten, bedarf es in der Europäischen Union der expliziten Lizenzierung für die Public Domain, ein Umstand, der im Falle schlecht verfügbarer Lizenzinformation zu erheblichen Folgeproblemen führen kann.

3.3. Befund 2: Heterogene Rechtelage

Die Analyse offenbart eine hohe Heterogenität an Lizenztypen allerdings mit regionalen Unterschieden.

Publicdata.eu weist 50 verschiedene Lizenztypen aus, allerdings ist es mit 55.481 Datensätzen ein vergleichsweise kleines Open Government Data Portal. 39,9% der Datensätze weisen individuelle Lizenzbestimmungen aus. 35,3% stehen unter einer Creative Commons Lizenz zur Verfügung und 0,5% unter Open Data Commons.

Das US Portal data.gov hält in Summe 132.206 Datensätze öffentlich zugänglich und weist 10 Lizenztypen aus. Diese bestehen zur Gänze aus Creative Commons, die jedoch nur 0,4% der verfügbaren Datensätze ausmachen. Die verbleibenden 99,6% weisen, wie bereits oben erläutert, keine Lizenzinformation aus.

OpenCanada, mit 244.257 Datensätzen das größte Datenportal, macht seinen Bestand unter drei Lizenztypen in Form individueller Lizenzvereinbarungen zugänglich.

Das Open Science Portal datahub.io weist 9.371 Datensätze und 33 Lizenztypen aus. Davon entfallen 19% auf individuelle Lizenzvereinbarungen, 17,1% auf Creative Commons und 4,8% auf Open Data Commons. Beachtliche 59,1% verfügen über keine Lizenzinformation.

Die hier skizzierten Unterschiede in der Lizenzgebarung lassen sich bisweilen auf Unterschiede in der Licensing Policy des jeweiligen Portals zurückführen. Die hohe Vielfalt an Lizenztypen in der europäischen Union ist auf den Umstand zurückzuführen, dass publicdata.eu eine freie Lizenzwahl erlaubt und jedes Mitgliedsland bis hin zu einzelnen Behörden individuelle Lizenzmodelle entwickelt hat. Dies trifft bisweilen auch auf datahub.io zu, das eine ähnliche Licensing Policy verfolgt.

In Kontrast dazu findet sich auf data.gov und OpenCanada nur eine beschränkte Zahl an Lizenztypen. Im Falle von OpenCanada kommen gar nur drei Lizenztypen zum Einsatz. Dies ist das Resultat der kanadischen Open

Data Directive¹⁴, welche vorschreibt, dass bei der Veröffentlichung von Daten auf OpenCanada zwingend zwischen drei Standardlizenzen auszuwählen ist. Dies hat auch zur Folge, dass 100% aller Datensätze mit Lizenzinformation ausgestattet sind. Erläuterungen der Rechtesituation in den USA finden sich im vorhergehenden Absatz.

3.4. Befund 3: Kompatibilitätskonflikte

Eine heterogene Rechtelage erhöht die Wahrscheinlichkeit von Lizenzkonflikten, welche die Kombination von unterschiedlichen Datensätzen ex-ante ausschließen. In der vorliegenden Untersuchung konnten Kompatibilitätskonflikte aufgrund der hohen Fallzahl nur stichprobenartig analysiert werden. Innerhalb der einzelnen Datenportale wurden unter den mit Lizenzen ausgewiesenen Datensätzen nur sehr geringe Anteile an Lizenzkonflikten diagnostiziert. Gemessen an der Open Definition¹⁵, d.h. der uneingeschränkten Offenlegung von Daten zum Wohle der Public Domain, liegt der Anteil an konfligierenden Lizenzen auf publicdata.eu bei 2%, auf data.gov bei 3% und auf datahub.io bei 7%.

Diese Werte muten auf den ersten Blick sehr niedrig an, dürfen jedoch nicht von dem Umstand ablenken, dass überwiegende Teile der verfügbaren Datensätze keine ausreichende Lizenzinformation zur Verfügung stellen, um etwaige Lizenzkonflikte überhaupt zu prüfen. Ausserdem ist zu beachten, dass der hohe Anteil an individuellen Lizenzmodellen als Folge der freien Lizenzwahl definitorische Spielräume geöffnet hat und im Zweifelsfall Raum für Interpretationskonflikte eröffnet. Eine strengere Lizenz überschreibt nicht zwangsläufig offenere Lizenzen und ist nicht zwingenderweise der kleinste gemeinsame Nenner eines derivativen Werkes, zumal Lizenzgeber festlegen können, dass eine Modifikation der Lizenzbedingungen nicht erlaubt ist. Einschränkungen in der Wiederverwendung ergeben sich bei Creative Commons durch die Bedingungen «non-derivative» (ND), «non-commercial» (NC) und «share-alike» (SA).¹⁶ Insbesondere die letzte Bedingung besagt, dass ein Datensatz seine Lizenzbedingungen auf einen neuen Datensatz überträgt und dadurch keiner strengeren Lizenz unterworfen werden darf. Komplexitätssteigernd wirkt auch der Copyleft-Passus, bei dem ein Datensatz seine Lizenzbedingungen auf das derivative Werk überträgt und damit die ursprüngliche Lizenz aushebelt.

Eine detaillierte Analyse der Kompatibilitätsproblematik steht zum aktuellen Zeitpunkt noch aus.

3.5. Befund 4: Maschinelle Erschließung von Lizenzinformation

Die Analyse der vier Datenportale offenbart weiters, dass zum Zeitpunkt der Analyse nur ein vernachlässigbarer Anteil der Datensätze seine Lizenzinformation in maschinen-lesbarer Form verfügbar macht. Wenige Ausnahmen existieren bei Datensätzen, die unter Creative Commons Lizenzen vertrieben werden und über einen Link auf eine maschinen-verarbeitbare Version der Lizenzbedingungen verweisen. Im Falle von publicdata.eu machen 2,6% der Datensätze von dieser Option Gebrauch. Im Falle von datahub.io sind es 2,2%.

Mit der zunehmenden Menge an Datensätzen und den Möglichkeiten der automatischen Syndizierung steigt auch der Bedarf nach der automatischen Klärung der Nutzungsrechte. Eine zentrale Rolle spielen hierbei sogenannte Lizenztechnologien und darauf aufbauende Filter- und Reasoning-Mechanismen, die auf Basis logischer Operatoren bei der Rechteklärung unterstützen.

¹⁴ Siehe <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=28108> (aufgerufen am 16. August 2015).

¹⁵ Siehe <http://opendefinition.org/> (aufgerufen am 16. August 2015).

¹⁶ Der Ausdruck «non-derivative» erlaubt keine Veränderung des Datensatzes. Der Ausdruck «non-commercial» verbietet die Verwendung des Datensatzes für kommerzielle Zwecke. Darunter versteht Creative Commons: «A commercial use is one primarily intended for commercial advantage or monetary compensation.» Siehe auch http://wiki.creativecommons.org/Frequently_Asked_Questions#Does_my_use_violate_the_NonCommercial_clause_of_the_licenses.3F (aufgerufen am 16. August 2015).

Zur maschinellen Auszeichnung von Licensing Policies wurden seit den 1990er Jahren sogenannte Rights Expression Languages (RELS) entwickelt, die dem Bereich der Digital Rights Management Technologien zuzurechnen sind [PRENAFETA 2011]. RELS unterstützen die Identifikation, Filterung, Syndizierung und Modifikation von Content, der sich aus mehreren Quellen unterschiedlicher Rechteinhaber speist, und sie bilden die Grundlage für eine differenzierte automatische Prozessierung von Content z.B. im Rahmen von eContracting- und eProcurement-Maßnahmen. RELS, wie etwa die Open Definitions Rights Language (ODRL)¹⁷ oder die Creative Commons Rights Expression Language (CCREL)¹⁸ sind damit eine zentrale technologische Komponente in service-orientierten, hoch automatisierten, vernetzten Verwertungsstrukturen.

Lizenztechnologien können einen wesentlichen Beitrag zur Herstellung von Rechtssicherheit in der Nutzung offener Daten leisten und gleichzeitig als nötige Basistechnologie für die Bewirtschaftung vernetzter Daten gesehen werden.

4. Conclusio und Ausblick

Die Verfügbarmachung und der automatische Austausch von Daten mittels technischer Schnittstellen zu kommerziellen und nicht-kommerziellen Zwecken stellt das Rechte- und Lizenzmanagement insbesondere ob der Fülle und Heterogenität an zugänglichen Datenquellen vor neue Herausforderungen. Noch hat sich keine Konvention zur Deklaration von Policies, die eine einfache Rechteklärung erlauben, ausgebildet. Die Situation wird durch regional und kulturell abweichende Rechtsregime und Lizenzierungspraktiken kompliziert. Auch hat sich die Verwendung von Rights Expression Languages bis dato nicht etabliert, wodurch nur eingeschränkte Möglichkeiten existieren, Datensätze auf Basis ihrer maschinenlesbaren Lizenzinformation einer automatischen Prozessierung zuzuführen.

Doch unter Bedingungen der digitalen Content-Bewirtschaftung liegt der Schlüssel zur gewerblichen Diversifikation in einer strategischen Diversifikation des Rechte- und Lizenzmanagements. Die notwendige Kulturtechnik im Sinne der maschinellen Bereitstellung von interoperablen Lizenzinformationen mittels Rights Expression Languages entlang des Urheber- und Datenbankrechts ist allerdings noch schwach ausgeprägt, was zum einen auf eine fehlende technische Infrastruktur in Form leicht bedienbarer, systemisch integrierter Tools zur Kompilation und Annotation von Datensätzen mit maschinenlesbaren Lizenzen, zum anderen auf fehlende ökonomische Incentives zur Veröffentlichung von Daten zurückzuführen ist. Vor dem Hintergrund der weiteren Ausdifferenzierung von datenbasierten Geschäftsmodellen ist jedoch von der inkrementellen Herausbildung einer Kulturtechnik der Datenlizenzierung auszugehen. Die technologischen und juristischen Voraussetzungen dafür sind geschaffen.

5. Literatur

ARCHER, PHIL; DEKKERS, MAX; GOEDERTIER, STIJN; LOUTAS, NIKOLAOS, Study on business models for Linked Open Government Data (BM4LOGD – SC6DI06692). Prepared for the ISA programme by PwC EU Services, 2013; Siehe http://ec.europa.eu/isa/documents/study-on-business-models-open-government_en.pdf (aufgerufen am 16. August 2015).

BACKHAUS, MARTIN, IT-gestütztes Rechte- und Lizenzmanagement im Rundfunk. Berlin: Logos, 2008.

FROSTERUS, MATTIAS; HYVONEN, EERO; LAITIO, JOONAS, Creating and publishing semantic metadata about linked and open datasets. In: Wood, David (Ed.). Linking Government Data. New York: Springer, 2011, S. 95–112.

GARCIA, ROBERTO; GIL, ROSA, Copyright licenses reasoning using an owl-dl ontology. In: Law, Ontologies and the Semantic Web: Channelling the Legal Information Flood. Amsterdam: IOS Press, 2009, S. 145–162.

GROVES, PETER, Sourcebook on Intellectual Property Law. London: Cavendish Publishing, 1997.

¹⁷ Siehe <http://www.w3.org/community/odrl/> (aufgerufen am 2. Januar 2015).

¹⁸ Siehe <http://www.w3.org/Submission/ccREL/> (aufgerufen am 2. Januar 2015).

- GUIBAULT, LUCIE; ANGELOPOULOS, CHRISTINA, *Open Content Licensing: From Theory to Practice*. Amsterdam: University Press, 2011.
- HOFFMANN, A., SCHULZ, T., ZIRFAS, J., HOFFMANN, H., ROSSNAGEL, A., & LEIMEISTER, J. M., Legal Compatibility as a Characteristic of Sociotechnical Systems: Goals and Standardized Requirements. *Business & Information Systems Engineering*, 57(2), 103–113. <http://doi.org/10.1007/s12599-015-0373-5>, 2015.
- HYLAND, BERNADETTE; WOOD, DAVID, The joy of data – a cookbook for publishing linked government data on the web. In: Wood, david (Ed.). *Linking government data*. New York: Springer, 2011, S. 3–26.
- HYVONEN, EERO, *Publishing and Using Cultural Heritage Linked Data on the Semantic Web*. Synthesis Lectures on the Semantic Web. Morgan & Claypool Publishers, 2012.
- JAIN, PRATEEK; HITZLER, PASCAL; JANOWICZ, KRZYSZTOF; VENKATRAMANI, CHITRA, There's No Money in Linked Data, 2013. In: <http://knoesis.wright.edu/faculty/pascal/pub/nomoneylod.pdf> (aufgerufen am 18. Dezember 2013)
- JENTZSCH, ANJA, *Linked Open Data Cloud*. In: Pellegrini, Tassilo; Sack, Harald; Auer, Sören (2014). *Linked Enterprise Data*. Berlin: Springer Verlag, 2014, S. 209–220.
- MAZZIOTTI, GIUSEPPE, *EU Digital Copyright Law and the End-User*. Heidelberg: Springer, 2008.
- PELLEGRINI, TASSILO, *Semantic Metadata in the News Production Process. Achievements and Challenges*. In: Lugmayr, Artur et al. (Eds). *Proceeding of the 16th International Academic MindTrek Conference 2012*. ACM SIGMM, 2012. S. 125–133.
- PELLEGRINI, TASSILO, *Linked Data Licensing – Datenlizenzierung unter netzökonomischen Bedingungen*. In: Schweighofer, Erich; Kummer, Franz; Hötzendorfer, Walter (Hrsg.). *Transparenz. Tagungsband des 17. Internationalen Rechtsinformatik Symposium IRIS 2014*. Wien: Verlag der Österreichischen Computergesellschaft, 2014, S. 159–168.
- PRENAFETA, JAVIER, *Protecting Copyright Through Semantic Technology*. In: *Publishing Research Quarterly*, 26/4, 2010, S. 249–254.
- PUCELLA, R.; WEISSMAN, V., *A logic for reasoning about digital rights*. In: *Proceedings of the 15th IEEE Computer Security Foundations Workshop*, IEEE, 2002, S. 282–294.
- ROTOLO, ANTONINO; VILLATA, SERENA; GANDON, FABIEN, *A deontic logic semantics for licenses composition in the web of data*. In: *Proceedings of the 14th International Conference on Artificial Intelligence and Law*, ACM, 2013, S. 111–120.
- SCHAAL, ULRICH, *Das strategische Management von Contentrechten. Schlüsselherausforderungen für audiovisuelle Medienunternehmen*. Wiesbaden: VS Verlag, 2010.
- SHY, OZ, *The Economics of Network Industries*. Cambridge: Cambridge University Press, 2001.
- SONNTAG, MICHAEL, *Rechtsschutz für Ontologien*. In: Schweighofer, Erich; Liebwald, Doris; Drachsler, Matthias; Geist, Anton (Eds.). *e-Staat und e-Wirtschaft aus rechtlicher Sicht*. Stuttgart: Richard Boorberg Verlag, 2006, S. 418–425.
- STALLMAN, RICHARD; LESSIG, LAWRENCE; GAY, JOSHUA, *Free Software, Free Society: Selected Essays of Richard M. Stallman*. Free Software Foundation, 2006.
- VÄLIMÄKI, MIKKO, *The Rise of Open Source Licensing: A Challenge to the Use of Intellectual Property in the Software Industry*. Helsinki: Turre Publishing, 2005.
- VILLATA, SERENA; GANDON, FABIEN, *Licenses compatibility and composition in the web of data*. *CEUR Workshop Proceedings*, 2012. Siehe auch: http://ceur-ws.org/Vol-914/paper_23.pdf (aufgerufen 15. August 2015).