

DeepGreen – DeepGold: Open-Access-Transformation – Entwicklung und Perspektiven

DeepGreen – DeepGold: open access transformation – development and perspectives

Abstract

The DeepGreen project (<https://deepgreen.kobv.de/en/project-deepgreen/>) is developing an automated workflow to transfer scholarly publications from publishers to repositories for open access publishing. Since 2011 German Alliance licenses include an open access component which allows authors to make their articles publicly available through their institutional or subject-based repositories after a shortened embargo period. DeepGreen is funded by the German Research Foundation (1st funding period: 01.2016–12.2017; 2nd funding period: 08.2018–07.2020) and is a consortium of 6 institutions:

- Cooperative Library Network Berlin-Brandenburg
- Bavarian State Library
- Bavarian Library Network
- University Library of the Technische Universität Berlin
- University Library of Erlangen-Nuremberg
- Helmholtz Open Science Coordination Office at the GFZ German Research Centre for Geosciences, Potsdam

Keywords: alliance licenses, gold open access, green open access, open access transformation

Zusammenfassung

Das deutsche Projekt DeepGreen (<https://deepgreen.kobv.de/de/deepgreen/>) arbeitet an einer automatisierten Lösung, um Artikeldaten (Volltexte und Metadaten) von wissenschaftlichen Verlagen abzuholen und an Repositorien zu liefern, die diese Artikel wiederum Open Access veröffentlichen können. Ausgangspunkt für das Projekt sind die überregional geförderten deutschen Allianz-Lizenzen. Diese enthalten eine Open-Access-Komponente, die es Autor*innen oder deren Institution erlaubt, ihre Artikel nach einer verkürzten Embargofrist in einem Repository ihrer Wahl Open Access zugänglich zu machen. DeepGreen wird seit 2016 von der Deutschen Forschungsgemeinschaft gefördert (1. Förderphase: 01.2016–12.2017; 2. Förderphase: 08.2018–07.2020) und besteht aus einem Konsortium aus 6 Institutionen:

- Kooperativer Bibliotheksverbund Berlin-Brandenburg (KOBV)
- Bayerische Staatsbibliothek (BSB)
- Bibliotheksverbund Bayern (BVB)
- Universitätsbibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)
- Helmholtz Open Science Koordinationsbüro am Deutschen GeoForschungsZentrum Potsdam (GFZ)
- Universitätsbibliothek der Technischen Universität Berlin (TUBB)

Schlüsselwörter: Open Access, Allianz-Lizenzen, DeepGreen, Goldener Open-Access-Weg, Grüner Open-Access-Weg, Open-Access-Transformation

Beate Rusch¹
Julia Boltze¹
Thomas Dierkes²
Julia Alexandra Goltz-Fellgiebel¹
Hedda Staub³

1 Kooperativer Bibliotheksverbund Berlin-Brandenburg, Berlin, Deutschland

2 Kooperativer Bibliotheksverbund Berlin-Brandenburg, Deutschland

3 Bayerische Staatsbibliothek, München, Deutschland

1 Einleitung

Open Access hat in den vergangenen Jahren enorm an Bedeutung gewonnen. Förderorganisationen, politische Entscheider und schließlich Open-Access-Aktivist*innen aus den Wissenschaften formulieren immer häufiger mit Nachdruck konkrete Zielvorgaben, um so Open Access zum Standard des Wissenschaftlichen Publizierens zu erheben. Beispielsweise wurde in Berlin 2015 mit der Open-Access-Strategie [1] für wissenschaftliche Institutionen des Landes das Ziel gesetzt, bis 2020 60% der wissenschaftlichen Zeitschriftenartikel im grünen oder goldenen Open Access verfügbar zu machen. Das Land Brandenburg feilt gerade an einem ähnlichen Strategiepapier. National setzte die Bundesregierung mit den 2016 vom Bundesministerium für Bildung und Forschung (BMBF) veröffentlichten Empfehlungen zu Open Access in der Wissenschaft [2] ein deutliches Zeichen Richtung Open-Access-Transformation. Auf internationaler Ebene gehen die Bestrebungen noch weiter: So wurde im September 2018 die cOAlition S (<https://www.coalition-s.org/>) ins Leben gerufen, ein Konsortium internationaler Forschungsförderinstitutionen, die mit ihrem Plan S festgelegt haben, dass ab 2020 alle wissenschaftlichen Publikationen zu öffentlich geförderten Forschungsprojekten Open-Access-konform veröffentlicht werden müssen. Politische Richtlinien wie diese haben zur Folge, dass immer mehr Institutionen ihre eigenen Open-Access-Strategien entwickeln und dass dadurch immer mehr institutionelle Repositorien entstehen, weiterentwickelt werden und gleichzeitig innerhalb ihrer Institution an Bedeutung gewinnen. Das Projekt DeepGreen agiert genau in diesem wachsenden Open-Access-Umfeld. Es möchte Wissenschaftler*innen, Bibliotheken und wissenschaftliche Institutionen dabei unterstützen, den freien Zugang zu publiziertem Wissen zu erhöhen. Das Projekt entwickelt eine technische Infrastruktur, um rechtssicher wissenschaftliche Publikationen, die lizenzrechtlich Open Access veröffentlicht werden dürfen, automatisiert von beteiligten Verlagen abzuholen und an berechnete Repositorien zu verteilen.

Dieser Artikel gibt einen Einblick in das Projekt DeepGreen (<https://deepgreen.kobv.de/de/deepgreen/>), angefangen bei der Entstehungsidee, über die bereits in der ersten Förderphase (01.2016–12.2017) erzielten Ergebnisse, bis hin zu den Meilensteinen, die die Projektpartner noch bis zum Ende der zweiten Förderphase (08.2018–07.2020) erreichen möchten. Dabei werden sowohl die rechtlichen Rahmenbedingungen, unter denen das Projekt agiert, erläutert, als auch der technische Prototyp der Datendrehscheibe, über die die Publikationen verteilt werden, beschrieben. Abschließend wird die Übertragbarkeit auf andere Länder diskutiert.

2 Was ist DeepGreen?

Das Projekt DeepGreen (Abbildung 1) wird seit 2016 von der Deutschen Forschungsgemeinschaft (DFG, [https://](https://www.dfg.de/)

www.dfg.de/) gefördert. Die Deutsche Forschungsgemeinschaft ist eine Selbstverwaltungseinrichtung, die Forschungsprojekte finanziell unterstützt und die internationale wissenschaftliche Zusammenarbeit fördert. Jährlich stehen der DFG ca. 3,2 Milliarden Euro zur Verfügung [3]. Mitte des Jahres 2014 wurde im Rahmen des Programms „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS) [4] die Ausschreibung Open-Access-Transformation veröffentlicht, in der um Anträge zu innovativen Lösungen für den Open-Access-Transformationsprozess gebeten wurde [5]. DeepGreen, von den Gutachter*innen anfänglich als Risikoprojekt eingestuft, konnte als eines der geförderten Projekte der Ausschreibung zunächst in eine zweijährige Projektphase (01.2016–12.2017) starten, die durch einen Folgeantrag auf vier Jahre Gesamtprojektlaufzeit erweitert wurde (Projektphase 2: 08.2018–07.2020). Aus Sicht der DFG lag und liegt das Risiko in der Bereitschaft der Verlage an DeepGreen teilzunehmen und eben darauf ist das Projekt zwingend angewiesen.



Abbildung 1: DeepGreen-Logo

Seinen ersten Ansatzpunkt hat DeepGreen bei den über die DFG seit 2011 geförderten sogenannten Allianz-Lizenzen. Von 2004 bis 2010 gab es in Deutschland die sogenannten Nationallizenzen. Diese enthielten Lizenzen für ausgewählte elektronische Zeitschriftenpakete und wurden zu 100% von der DFG gefördert [6]. 2011 wurden sogenannte Allianz-Lizenzen als Nachfolgelizenzen eingeführt, die im Gegensatz zu den Nationallizenzen nur noch zu 25% von der DFG gefördert wurden. Die verbleibenden 75% mussten von jeder teilnehmenden Institution selbst getragen werden [7]. Zusätzlich zu den Bedingungen der Nationallizenzen wurde in den Allianz-Lizenzen mit den Verlagen eine Open-Access-Komponente verhandelt, die besagt:

„Autoren aus autorisierten Einrichtungen sind ohne Mehrkosten berechtigt, ihre in den lizenzierten Zeitschriften erschienenen Artikel in der Regel in der durch den Verlag publizierten Form (z.B. PDF) zeitnah in institutionelle oder disziplin-spezifische Repositorien ihrer Wahl einzupflegen und im Open Access zugänglich zu machen. Das gleiche Recht besitzen die autorisierten Einrichtungen, denen die jeweiligen Autoren angehören.“ ([8], S. 8)

Diese Open-Access-Komponente geht deutlich über das allgemein von den Verlagen eingeräumte Zweitveröffentlichungsrecht hinaus. Das betrifft sowohl die zur Zweitveröffentlichung freigegebene Version als auch die Embargofrist. Die Allianz-Lizenz-Regelung erlaubt jeder Institution, die Artikel ihrer eigenen Wissenschaftler*innen nach ei-

ner verkürzten Embargofrist als Verlags-PDF (einige der beteiligten Verlage erlauben nur die Verwendung des Autorenmanuskripts) Open Access in ihrem (institutionellen und/oder fachbezogenen) Repositorium zu veröffentlichen.

Obwohl durch vertragliche Komponenten wie diese eine Zweitveröffentlichung wissenschaftlicher Publikationen in Deutschland durchaus erlaubt ist und auch bezahlt wurde, werden diese Rechte nur selten in Anspruch genommen. Autor*innen sind unsicher und kennen die rechtlichen Grundlagen häufig nicht im Detail (zu beachten sind individuelle Autor*innenverträge, institutionsbezogene und nationale Lizenzverträge, sowie das seit 2014 für eine eingeschränkte Gruppe an Wissenschaftler*innen geltende deutsche Zweitveröffentlichungsrecht [9]), so dass viele Forschende zögern, ihre wissenschaftlichen Publikationen in ihr institutionelles Repositorium zu stellen ([10], S. 12). Zu dieser rechtlichen Unsicherheit kommt, dass die Autor*innen oftmals gar nicht wissen, dass ihre Institution Lizenzen abgeschlossen hat, die auch ihre Publikationen betreffen. Damit wird es zur Aufgabe der jeweiligen Institution, sich um die Wahrnehmung der Zweitveröffentlichungsrechte ihrer Forschenden zu kümmern und den Open-Access-Anteil in ihrem Repositorium zu erhöhen. Dieses Aufgabenfeld übernimmt in der Regel die angeschlossene Bibliothek.

Bei der Umsetzung ergeben sich in Bezug auf die Allianz-Lizenzen einige Herausforderungen: Zum einen muss jede teilnehmende Institution wissen, welche ihrer Autor*innen in welchem Allianz-Lizenzverlag publiziert und zum anderen müsste jede Institution die jeweiligen Verlage um die einzelnen Artikel bitten, zu deren Open-Access-Veröffentlichung sie berechtigt ist. Diese Daten zu ermitteln bindet personelle Ressourcen, die in der Regel bereits für Publikationsunterstützung und bibliothekarische Beratungsleistungen benötigt werden. Und die rechtlichen Schwierigkeiten und Bedenken hinsichtlich der Zweitveröffentlichungsrechte betreffen nicht nur die Autor*innen sondern werden auch von Bibliotheken als deren lizenzberechtigte Vertretung geteilt ([11], S.7). 2013 – zwei Jahre nach Abschluss der neuen Lizenzverträge – führte die Bayerische Staatsbibliothek als einer der Verhandlungsführer von Allianz-Lizenzen eine Umfrage zur Nutzung der verhandelten Open-Access-Rechte durch. Das Ergebnis war ernüchternd. Nur sehr wenige Institutionen nutzten die Open-Access-Komponente aktiv und auch seitens der Autor*innen bestand nur geringes Interesse ([12], S.1). DeepGreen will eine automatisierte Lösung bieten, um Institutionen und Wissenschaftler*innen dabei zu unterstützen, die Wahrnehmung ihrer Open-Access-Rechte leichter zu bewerkstelligen. Hier sieht das Projekt nicht nur die automatisierte Verteilung von Publikationen in Bezug auf die Allianz-Lizenzen, auch die Übertragbarkeit auf andere Lizenzarten wird diskutiert.

DeepGreen besteht aus einem Projektkonsortium aus sechs Institutionen: Die strategischen Bibliotheksverbundpartner Kooperativer Bibliotheksverbund Berlin-Brandenburg (KOBV) und Bibliotheksverbund Bayern (BVB) agieren gemeinsam mit einem der Verhandlungsführer für Allianz-

Lizenzen, der Bayerischen Staatsbibliothek (BSB), sowie drei lizenzteilnehmenden Einrichtungen – den Universitätsbibliotheken der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und der Technischen Universität Berlin (TUB) sowie dem Helmholtz Open Science Koordinationsbüro am Deutschen GeoForschungsZentrum (GFZ).

Jede Institution ist für die Bearbeitung verschiedener Teilbereiche innerhalb des Projektes verantwortlich, wobei neben der Projektkoordination die technische Entwicklung und Implementierung beim KOBV liegt.

2.1 Erste Projektphase, Januar 2016 – Dezember 2017

Innerhalb der ersten Projektphase wurde zunächst mit den vierzehn (Allianz-Lizenz-)Verlagen eruiert, welche Anforderungen an ein mögliches Metadatenschema für die technische Plattform bestehen [13]. In einem Workshop wurden anschließend gemeinsam mit Verlagen und Vertreter*innen institutioneller Repositorien Anforderungen an die von DeepGreen zu entwickelnde Datendrehzscheibe und den Gesamtworkflow diskutiert. Nach Abwägung zwischen einer kompletten Eigenentwicklung der benötigten technischen Infrastruktur oder aber der Nachnutzung bereits existierender Software aus ähnlichen Projekten fiel die Entscheidung schließlich auf die Verwendung der 2015 frei auf GitHub zur Verfügung gestellten Software des Publications Event Router (<https://pubrouter.jisc.ac.uk/>) des Joint Information Systems Committee (Jisc) aus Großbritannien.

Auf Basis dieser Open-Source-Software wurde innerhalb der ersten Projektphase der Prototyp der DeepGreen-Datendrehzscheibe entwickelt und implementiert. Durch Einigungen mit den Verlagen S. Karger und SAGE Publications wurde das vom Forschungsförderer geforderte Ziel erreicht, eine Kooperation mit mindestens zwei Verlagen aufzubauen. Zusätzlich wurden testweise Daten der Verlage Royal Society of Chemistry, Oxford University Press, European Mathematical Society und De Gruyter geliefert und über die DeepGreen-Datendrehzscheibe verarbeitet. Vor allem die Verlage S. Karger und SAGE Publications konzentrieren sich in ihrem Portfolio zum größten Teil auf medizinische Themenbereiche, was DeepGreen für medizinische und lebenswissenschaftliche Forschung attraktiv macht.

Als Abschluss der ersten Projektphase konnte der Prototyp der Datendrehzscheibe mit den Repositorien der Partnerinstitutionen FAU (eingesetzte Repositoriensoftware: OPUS 4), TU Berlin (mit einer DSpace-Instanz) und GFZ (mit einem EsciDoc/Pubman-Repositorium) erfolgreich getestet und auf einem öffentlichen Workshop interessierten Institutionen präsentiert werden.

2.2 Zweite Projektphase, August 2018 – Juli 2020

Im Frühjahr 2018 stimmte die DFG einer Verlängerung des Projektes zu, so dass DeepGreen im August 2018 in

eine weitere zweijährige Förderphase starten konnte. Diese Projektphase verfolgt zwei Ziele. Zum einen soll DeepGreen mit einem nachhaltigen Geschäfts- und Betriebsmodell als nationaler Dienst (mit Schwerpunkt auf Allianz-Lizenzen) etabliert werden. Zusätzlich will das Projektkonsortium in dieser Projektphase prüfen, welche weiteren Lizenzmodelle für DeepGreen in Frage kommen, beispielsweise die derzeit zu verhandelnden Lizenzen für Fachinformationsdienste oder unterschiedliche Konsortiallizenzen ohne explizite grüne Open-Access-Komponente. Auch soll geprüft werden, ob die DEAL-Verträge [14] einen möglichen Ansatzpunkt für DeepGreen darstellen. Neben S. Karger und SAGE Publications sollen weitere Verlage für eine Kooperation gewonnen werden. Zusätzlich zu institutionellen Repositorien sollen zukünftig auch Fachrepositorien bedient werden, wofür die Entwicklung neuer Zuordnungsworkflows benötigt wird. Außerdem soll die Einbindung goldener Open-Access-Verlage möglich werden. Parallel zu den genannten Zielen erfolgt eine stetige Konsolidierung und Weiterentwicklung der DeepGreen-Datendrehscheibe.

3 Vertragliche Voraussetzungen

Zur Wahrnehmung vorhandener Open-Access-Rechte durch einen maschinellen Service wie DeepGreen (so wie er momentan implementiert ist) sind zunächst die notwendigen vertraglichen Voraussetzungen zu identifizieren:

1. Der Vertrag wurde in einem Konsortium für mehrere Bibliotheken ausgehandelt, da nur konsortiale Open-Access-Regelungen in der Elektronischen Zeitschriftenbibliothek nachgewiesen sind. Diese Informationen sind für die Zuordnung grüner Open-Access-Artikel durch DeepGreen zwingend notwendig.
2. Der Vertrag enthält Zweitveröffentlichungsrechte für Angehörige der lizenznehmenden Institutionen, die auch von den betreffenden Institutionen selbst wahrgenommen werden dürfen.
3. Die Zweitveröffentlichung darf in institutionellen und/oder fachspezifischen Repositorien gespeichert und dort im Open Access frei zur Verfügung gestellt werden.
4. Optional kann auch die Lieferung der betroffenen Publikationen und ihrer Metadaten durch den Verlag (über DeepGreen) vereinbart werden.

Die Open-Access-Komponenten sind sowohl in der Elektronischen Zeitschriftenbibliothek (EZB) als auch in einer Übersicht für die DFG-geförderten Lizenzen für elektronische Medien verzeichnet [15]. Die EZB ist ein Angebot der Universität Regensburg zur Nutzung elektronischer wissenschaftlicher Zeitschriften. Dabei werden nicht nur die Zugriffsberechtigungen einzelner Institutionen auf verschiedene Zeitschriften abgebildet, sondern es wird auch verzeichnet, unter welcher Lizenz diese Berechtigungen stehen [16]. Für die Nutzung von DeepGreen ist eine Teilnahme an der EZB daher momentan verpflichtend.

In Deutschland erfüllen neben Gold-Open-Access-Vereinbarungen v.a. die Allianz-Lizenzen, aber auch einige Nationallizenzen die oben genannten Voraussetzungen, so dass die berechtigten Institutionen DeepGreen den Auftrag zu einer maschinellen Verarbeitung und damit der effizienten Wahrnehmung ihrer Rechte geben können.

Eine Verarbeitung durch DeepGreen ist bisher in keinem dieser Verträge festgeschrieben, so dass zusätzliche Verhandlungen zwischen DeepGreen und dem jeweiligen Verlag notwendig sind, um eine Datenlieferung bzw. Datenverteilung via DeepGreen aufzubauen. Neben der Auswertung von Lizenzverträgen und einem Auftrag der teilnehmenden Institutionen ist auch eine direkte Kooperation zwischen DeepGreen und reinen Open-Access-Verlagen möglich, da die entsprechenden Publikationsverträge ganz grundsätzlich eine Weiterverwendung der Publikationen erlauben.

4 Technische Umsetzung der DeepGreen-Datendrehscheibe

DeepGreen funktioniert wie ein Postamt: Verlage schicken Daten, die dann von DeepGreen gemäß den rechtlichen Bedingungen an die jeweils berechtigten Empfänger (einer Institution zugehörige Repositorien als Vertretung für die Autor*innen) zugestellt werden.

Als Programmierplattform für dieses Online-System wird das Micro-Webframework Flask (<http://flask.pocoo.org/>, <https://github.com/pallets/flask>) unter Python verwendet. Mit Flask lassen sich sehr einfach in kurzer Zeit serverbasierte Applikationen erstellen, die mit einem beliebigen Web-Browser aufgerufen werden können. Die Bereitstellung der einzelnen Funktionalitäten erfolgt dabei über sogenannte REST-APIs (engl. representational state transfer application programming interface). Die gesamte Datenverwaltung übernimmt Elasticsearch [17], ein NoSQL-Datenbanksystem, welches auf der Basis von Lucene [18], einer Datenbankbibliothek zur Volltextsuche von Apache, entwickelt wurde.

4.1 Kontobasierter Service

Sowohl Verlagshäuser als auch die einzelnen Repositorien bekommen in der Datendrehscheibe eigenständige Konten zugewiesen, über die der gesamte Datenverkehr abgewickelt wird. Über die Verlagskonten können über geeignete Schnittstellen Publikationsmeldungen in Form von Metadaten (ggf. gemeinsam mit den zugehörigen Volltexten) an DeepGreen geschickt werden. Die Repositorienkonten erhalten Metadaten (und die dazugehörigen Volltexte) ebenfalls über geeignete Schnittstellen (z.B. auch über http[s]) bereitgestellt, so sie lizenzbedingt für eine weitere Nutzung der Artikel berechtigt sind.

Jedes Konto ist an einen sogenannten, nur dem Konto bekannten API-Key in Form eines 32-stelligen Hexadezimalcodes gekoppelt. Mit diesem Schlüssel wird der Zugang zu den Schnittstellen von DeepGreen gewährt. So wird eine effektive Zugriffsberechtigung auf die ent-

sprechenden Publikationsdaten gewährleistet. Falls nötig kann dieser Schlüssel erneuert werden.

4.2 Schnittstellen von DeepGreen

Die Datenübertragung rund um DeepGreen ist über mehrere Schnittstellen möglich: Verlage können über sFTP, SWORD (v2) oder über die DeepGreen-REST-API Metadaten und Volltexte an den DeepGreen-Server abliefern. Für den Import in Repositorien stehen OAI-PMH, SWORD (v2) oder auch die DeepGreen-eigene REST-API zur Verfügung. Über OAI-PMH und die DeepGreen REST-API werden zunächst nur die Metadaten einer OA-fähigen Publikation an die berechtigten Repositorien gesendet, die dazugehörigen Volltexte müssen im Anschluss als zip-Paket über eine geheime, d.h. eine https-Adresse mit dem, nur dem jeweiligen Repositorium bekannten API-Key, bezogen werden. Die genauen Funktionsweisen von OAI-PMH [19] und sFTP [20] sind ausführlich dokumentiert [21], [22]. Die REST-API von DeepGreen (mit der Basisadresse <https://www.aa-deepgreen.de> für die GET- und POST-Befehle) ist wie folgt strukturiert:

- **Verlage**
 POST /api/v1/validate?api_key<api_key>
 POST /api/v1/notification?api_key=<api_key>
- **Repositorien**
 GET /api/v1/notification/<note_id>
 GET /api/v1/notification/<note_id>/content[/fname]
 ?api_key=<api_key>
 GET|POST /api/v1/config/<repo_id>?api_key=
 <api_key>
- **Verlage & Repositorien**
 GET /api/v1/routed?<parameter>
 GET /api/v1/routed/<repo_id>?<parameter>

<parameter>: since=<YYYY-MM-DD>, pageSize=<N> oder page=<N>

Als Ein- und Ausgabeformate werden verschiedene XML-basierte Schemata verwendet, z.B. NISO Journal Article Tag Suite (JATS) [23] für die Datenlieferung durch die Verlage und METS/MODS, OPUS-XML oder Dublin Core für die Metadatenabgabe an berechnigte Repositorien. Die Unterstützung weiterer Formate kann entwickelt werden.

4.3 Mechanismus der Lizenzüberprüfung

Die Zuordnungslogik, die entscheidet, ob ein Artikel einer Einrichtung zur weiteren Verarbeitung zur Verfügung gestellt werden darf oder nicht, erfolgt in zwei Schritten. Zunächst wird aus den Metadaten eines jeden Artikels die ISSN der Zeitschrift ermittelt, anhand derer geprüft wird, ob der Artikel aus einer Zeitschrift stammt, für die eine OA-fähige Lizenz vorliegt. Dieser Abgleich erfolgt in systemseitig gespeicherten Listen, die zwei Informationen aus der EZB vereinen: welcher Verlag mit welcher Zeitschrift an einem Lizenzvertrag teilnimmt und welche (deutschsprachigen) Einrichtungen an den OA-fähigen

Lizenzen beteiligt und somit berechnigt sind, die Werke ihrer Autor*innen zu erhalten. Dieser Schritt wird für reine Open-Access-Verlage übersprungen.

In einem zweiten Schritt findet die eigentliche Zuordnung statt, indem geprüft wird, ob in den Angaben der Affiliationen der Autoren des Artikels (die wiederum aus den vorliegenden Metadaten gewonnen werden) Namensansetzungen einer zuvor als berechnigt erkannten Einrichtung (oder von mehreren als berechnigt erkannten Einrichtungen) vorkommen. Ist dies der Fall, bekommen die betreffenden Einrichtungen den analysierten Artikel zugewiesen und können den Volltext und die Metadaten gemäß der festgestellten Lizenz weiterverarbeiten. Jede vorgenommene Zuordnung wird intern protokolliert und dem betroffenen Verlag mitgeteilt, jeder Artikel der keiner Institution zugeordnet werden konnte, wird ebenso dokumentiert. Dies ermöglicht den Verlagen die vollständige und transparente Kontrolle über die Verarbeitung ihrer Publikationen.

Die Namensansetzungen bzw. -varianten einer wissenschaftlichen Einrichtung und weitere, eindeutige Angaben in einer Publikation mit Institutionsbezug verwaltet jede Einrichtung über ihr Konto selbsttätig. Für jedes Repositorium besteht die Möglichkeit, die Angaben zur Affiliation als einfache Textdatei in das eigene Konto hochzuladen. Weitere Angaben zu Affiliationen umfassen z. B. Förderkürzel, wie sie in Veröffentlichungen anzutreffen sind, oder auch Domainnamen (der zweite Teil einer E-Mail-Adresse nach dem @-Zeichen).

Die interne Datenstruktur von DeepGreen ist so konzipiert, dass verschiedene Lizenzarten abgebildet werden können. So liegen DeepGreen bisher Pakete/Kollektionen von Allianz- und Nationallizenzen einzelner Verlage sowie Pakete von Gold-Open-Access-Verlagen vor. Für die Gold-Open-Access-Publikationen verzichtet der Zuordnungsmechanismus auf die Überprüfung der Teilhabe an einer Allianz-Lizenz und prüft sofort die Affiliation für alle vorliegenden Repositorienkonten. Die Angaben zu Allianz- bzw. Nationallizenzen bezieht DeepGreen periodisch über die Elektronische Zeitschriftenbibliothek (EZB) in Regensburg.

5 Übertragbarkeit auf andere Länder

International besteht großes Interesse an einem Service wie DeepGreen, denn viele Repositorienbetreiber wünschen sich eine automatisierte Umsetzung grüner und goldener Open-Access-Rechte. Doch ist DeepGreen überhaupt für andere Länder denkbar?

Die Beantwortung dieser Frage hängt stark davon ab, wie die Repositorienlandschaft des jeweiligen Landes aussieht. Bei einer sehr zersplitterten und heterogenen Repositorieninfrastruktur wie in Deutschland, wo fast jede Forschungsinstitution ihr eigenes Repositorium betreibt, wäre es nötig, in Analogie zu DeepGreen eine neue technische Infrastruktur aufzubauen. Ist aber, wie beispielsweise in Frankreich, bereits ein etabliertes nationales Repositorium vorhanden, würde es ausreichen, ent-

sprechende Datenlieferungen mit den Verlagen zu verhandeln. Diese Verhandlungen würden umso einfacher, je selbstverständlicher diese Lieferungen von allen Ländern gefordert würden. Würden Services wie DeepGreen weltweit kooperieren und ein internationales Netzwerk bilden, wäre es überdies leichter, nicht nur einheitliche Qualitätsstandards für Metadaten in der Praxis zu etablieren, sondern auch bereits existierende Standards wie ORCID zu stärken.

Seit Anfang des Jahres 2019 gibt es erste Gespräche zwischen dem DeepGreen-Projektkonsortium und dem österreichischen Netzwerk für RepositorienmanagerInnen über eine mögliche Kooperation, wobei die Idee, Artikel automatisiert von Verlagen abzuholen und an berechnete Repositorien weiterzuleiten, mit großem Interesse aufgenommen wurde. Die Umsetzung eines Service wie DeepGreen bzw. eine direkte Kooperation mit den deutschen DeepGreen-Partnern wird seitdem für die österreichische Repositorienlandschaft diskutiert. Denkbar ist sowohl der Aufbau eines eigenen österreichischen „DeepGreen-Knotens“ an einer geeigneten Einrichtung wie auch die Versorgung der österreichischen Universitäten und Hochschulen durch den „deutschen Knoten“ mit den entsprechenden Anpassungen – solche wären vermutlich hinsichtlich des Ausgabeformates für die Repositoriensoftware Visual Library (https://www.semantics.de/visual_library/), die in Österreich weit verbreitet ist, notwendig.

Neben dem technischen Betrieb ist eine vertrauensvolle Kooperation und Kommunikation mit Verlagen wie auch mit den Wissenschaftseinrichtungen zwingende Voraussetzung für das Gelingen von DeepGreen. Selbst bei einer Übernahme des technischen Parts von DeepGreen durch einen deutschen Partner bliebe damit für die Einrichtungen vor Ort die Aufgabe, die vertraglichen Voraussetzungen mit den Verlagen zu klären und als Ansprechpartner zur Verfügung zu stehen.

Sollen vor allem goldene Open-Access-Komponenten in konsortialen Verträgen mit einem DeepGold-Knoten umgesetzt werden, liegt der Mehrwert einer solchen Plattform darin, dass die Publikationen zusätzlich zum verlags-eigenen Angebot automatisiert an institutionelle und/oder fachspezifische Repositorien gemeldet und dann dort zusätzlich vorgehalten und zugänglich gemacht werden können, womit auch dem Gedanken der Archivierung wissenschaftlichen Outputs unabhängig von möglichen Verlagsangeboten Rechnung getragen wäre. Durch diese zusätzliche Erschließung wird also nicht nur die Auffindbarkeit und Sichtbarkeit der Artikel, sondern auch die Vollständigkeit der Sammlung des jeweiligen Repositoriums vorangetrieben.

6 Ausblick

Im Sommer 2019 startet das Projekt DeepGreen mit knapp 30 institutionellen Repositorien aus der ganzen Bundesrepublik Deutschland in eine erweiterte Testphase. In Vorbereitung darauf wurde eine Auswahl institutio-

neller Repositorien kontaktiert, die Allianz-Lizenzen mit den Verlagen S. Karger und/oder SAGE Publications besitzen. Der Fokus für die erweiterte Testphase liegt vor allem auf den Allianz-Lizenzen, jedoch ist auch geplant, prototypisch die Verteilung von Artikeln in den Lizenzkontexten „Gold Open Access“ und „Nationallizenz“ zu testen. Diese breit angelegte Testphase soll genutzt werden, um Erfahrungen in Bezug auf großvolumige Datenlieferungen durch die Verlage, die Anbindung und den Umgang mit diversen Softwaretypen der datenabnehmenden Repositorien (DSpace, Eprints, EsciDoc/Pubman, InVenio, MyCore, OPUS 4, etc.) und den nötigen Umfang eines Supports für die am Service teilnehmenden Repositorienbetreiber zu sammeln. Im Anschluss folgt eine strukturierte Feedbackrunde, die dazu dient, Einblick in die Erfahrungen der Repositorienbetreiber zu erhalten und mögliche weitere Mehrwerte des Services zu diskutieren. Die erfolgreiche erweiterte Testphase soll auch zu Kooperationen mit weiteren Verlagen führen.

Ziel ist es, DeepGreen zu einem bundesweiten Service zu entwickeln. In diesem Zusammenhang muss ein nachhaltiges Betriebs- und Finanzierungsmodell entwickelt werden. Die deutsche Bibliothekslandschaft ist in sechs Bibliotheksverbände aufgeteilt, in denen der überwiegende Teil der potentiellen DeepGreen-Teilnehmer organisiert ist. Eine mögliche Idee für ein Finanzierungsmodell wäre, die Kosten von DeepGreen auf die Bibliotheksverbände aufzuteilen, die ihren Teilnehmerbibliotheken wiederum den DeepGreen-Service als Verbunddienstleistung anbieten könnten. Bis Ende 2019 soll die Konzeption eines realisierbaren und pragmatischen Betriebs- und Finanzierungsmodells abgeschlossen sein.

7 Veröffentlichungen zu DeepGreen

Für nähere Informationen zu DeepGreen möchten wir auf die Veröffentlichungen zum Projekt hinweisen:

- Projektanträge [12], [24]
- Zeitschriftenartikel [25], [26], [27]
- Blogbeitrag [28]

Anmerkung

Interessenkonflikte

Die Autoren erklären, dass sie keine Interessenkonflikte in Zusammenhang mit diesem Artikel haben.

Förderung

Diese Arbeit wurde unter dem Förderkennzeichen (grant number) [401038572] durch die Deutsche Forschungsgemeinschaft unterstützt.

Literatur

1. Der Senat von Berlin. Open-Access-Strategie für Berlin: wissenschaftliche Publikationen für jedermann zugänglich und nutzbar machen. 2015. Verfügbar unter: <https://www.parlament-berlin.de/adoss/17/IIIPlen/vorgang/d17-2512.pdf>
2. Bundesministerium für Bildung und Forschung. Open Access in Deutschland. Die Strategie des Bundesministeriums für Bildung und Forschung. 2016. Verfügbar unter: https://www.bmbf.de/upload_filestore/pub/Open_Access_in_Deutschland.pdf
3. Deutsche Forschungsgemeinschaft. Was ist die DFG? Verfügbar unter: https://www.dfg.de/dfg_profil/aufgaben/was_ist_die_dfg/index.html
4. Deutsche Forschungsgemeinschaft. Wissenschaftliche Literaturversorgungs- und Informationssysteme. Verfügbar unter: <https://www.dfg.de/foerderung/programme/infrastruktur/lis/index.html>
5. Deutsche Forschungsgemeinschaft. Open-Access-Transformation. Verfügbar unter: https://www.dfg.de/foerderung/info_wissenschaft/2014/info_wissenschaft_14_29/index.html
6. Deutsche Forschungsgemeinschaft. Nationallizenzen „Classics“. Verfügbar unter: <https://www.nationallizenzen.de/ueber-nationalizenzen/nationallizenzen#1.6>
7. Deutsche Forschungsgemeinschaft. Allianz-Lizenzen. Verfügbar unter: <https://www.nationallizenzen.de/ueber-nationalizenzen/allianz-lizenzen-2011-ff>
8. Deutsche Forschungsgemeinschaft. Grundsätze für den Erwerb DFG-geförderter überregionaler Lizenzen. Verfügbar unter: <https://www.nationallizenzen.de/ueber-nationalizenzen/al-grundsätze>
9. Pampel H. Zweitveröffentlichungsrecht: Die Richtung stimmt, die Details enttäuschen. 2013. Verfügbar unter: <https://irights.info/artikel/zweitveröffentlichungsrecht-die-richtung-stimmt-die-details-enttäuschen/15422>
10. Morais R, Borrell-Damian L. Open Access 2016-2017 EUA Survey Results. Brussels: European University Association; 2018. Verfügbar unter: <https://eua.eu/resources/publications/324:open-access-in-european-universities-results-from-the-2016-2017-eua-institutional-survey.html>
11. Blasetti A, Golda S, Göhring D, Grimm S, Kroll N, Sievers D, Voigt M. Smash the Paywalls: Workflows und Werkzeuge für den grünen Weg des Open Access. Informationspraxis. 2019;5(1). Verfügbar unter: <https://journals.ub.uni-heidelberg.de/index.php/ip/article/view/52671>
12. Christof J, Bertelmann R, Griebel R, Koch T, Söllner K. DeepGreen – Entwicklung eines rechtssicheren Workflows zur effizienten Umsetzung der Open-Access-Komponente in den Allianz-Lizenzen für die Wissenschaft. ZIB-Report15-58. 2016. URN: urn:nbn:de:0297-zib-56799
13. Becker P, Bertelmann R, Ceynowa K, Christof J, Dierkes T, Goltz JA, et al. Questionnaire for effective exchange of bibliographic metadata – current status of publishing houses. 2016. URN: urn:nbn:de:0297-zib-60419
14. Projekt DEAL. Verfügbar unter: <https://www.projekt-deal.de/>
15. Deutsche Forschungsgemeinschaft. Nationallizenzen. Open Access. Verfügbar unter: <https://www.nationallizenzen.de/open-access>
16. Universität Regensburg. Elektronische Zeitschriftenbibliothek. Verfügbar unter: <https://rzblx1.uni-regensburg.de/>
17. Elasticsearch. In: Wikipedia. Verfügbar unter: <https://de.wikipedia.org/wiki/Elasticsearch>
18. Apache Lucene. In: Wikipedia. Verfügbar unter: https://de.wikipedia.org/wiki/Apache_Lucene
19. Wikipedia. Open Archives Initiatives. In: Wikipedia. Verfügbar unter: https://de.wikipedia.org/wiki/Open_Archives_Initiative
20. SSH File Transfer Protocol. In: Wikipedia. Verfügbar unter: https://de.wikipedia.org/wiki/SSH_File_Transfer_Protocol
21. Open Archives Initiative. The Open Archives Initiative Protocol for Metadata Harvesting. Document Version 2015-01-08. Verfügbar unter: <http://www.openarchives.org/OAI/openarchivesprotocol.html>
22. Galbraith J, Saarenmaa O. SSH File Transfer Protocol. Internet-Draft. 2006. Verfügbar unter: <https://tools.ietf.org/html/draft-ietf-secsh-filexfer-13>
23. National Information Standards Organization (NISO). JATS: Journal Article Tag Suite. Verfügbar unter: <http://jats.niso.org/1.1/>
24. Koch T, Ceynowa K, Söllner K, Christof J, Bertelmann R. DeepGreen – Open Access Transformation. Etablierung und Weiterentwicklung rechtssicherer Workflows zur effizienten Umsetzung von Open-Access-Komponenten in Lizenzvereinbarungen für wissenschaftliche Publikationen. ZIB-Report 18-39. 2018. URN: urn:nbn:de:0297-zib-69612
25. Putnings M, Rusch B. DeepGreen – Entwicklung eines rechtssicheren Workflows zur effizienten Umsetzung der Open-Access-Komponente in den Allianz-Lizenzen für die Wissenschaft. o-bib – das offene Bibliotheksjournal. 2016;3(4):110-9. DOI: 10.5282/o-bib/2016H4S110-119
26. Dierkes T. Ist die Umsetzung der Open-Access-Komponente in Allianz-Lizenzverträgen wirklich so kompliziert? ABI Technik. 2017;37(3):237. DOI: 10.1515/abitech-2017-0051
27. Goltz-Fellgiebel JA, Putnings M. Open-Access-Transformation mit DeepGreen: Gemeinsam den (grünen) Schatz heben. o-bib – das offene Bibliotheksjournal. 2019;6(1):1-11. DOI: 10.5282/o-bib/2019H1S1-11
28. Scheliga K, Goltz J. DeepGreen: Open-Access-Transformation. ZBW MediaTalk. 2017. Verfügbar unter: <https://www.zbw-mediatalk.eu/2017/06/deepgreen-open-access-transformation/>

Korrespondenzadresse:

Beate Rusch
Kooperativer Bibliotheksverbund Berlin-Brandenburg,
Takustr. 7, 14195 Berlin, Deutschland
rusch@zib.de

Bitte zitieren als

Rusch B, Boltze J, Dierkes T, Goltz-Fellgiebel JA, Staub H. DeepGreen – DeepGold: Open-Access-Transformation – Entwicklung und Perspektiven. GMS Med Bibl Inf. 2019;19(1-2):Doc07. DOI: 10.3205/mbi000432, URN: urn:nbn:de:0183-mbi0004322

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/mbi/2019-19/mbi000432.shtml>

Veröffentlicht: 10.09.2019

Copyright

©2019 Rusch et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.