

DeepGreen

Eine Infrastruktur für die Open-Access-Transformation

Julia Boltze, Zuse Institut Berlin, Kooperativer Bibliotheksverbund Berlin-Brandenburg

Annika Höllerl, Bayerische Staatsbibliothek München

Monika Kuberek, Universitätsbibliothek der Technischen Universität Berlin

Stefan Lohrum, Zuse Institut Berlin, Kooperativer Bibliotheksverbund Berlin-Brandenburg

Heinz Pampel, Helmholtz-Gemeinschaft, Helmholtz Open Science Office, Potsdam

Markus Putnigs, Universitätsbibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg

Regina Retter, Bayerische Staatsbibliothek München

Beate Rusch, Zuse Institut Berlin, Kooperativer Bibliotheksverbund Berlin-Brandenburg

Hildgard Schäffler, Bayerische Staatsbibliothek München

Konstanze Söllner, Universitätsbibliothek der Friedrich-Alexander-Universität Erlangen-Nürnberg

Zusammenfassung

Mit DeepGreen wurde eine Infrastruktur aufgebaut und etabliert, die Zeitschriftenartikel von wissenschaftlichen Verlagen abholt und berechtigten Bibliotheken zur Veröffentlichung in ihren Repositorien sendet. DeepGreen unterstützt Bibliotheken als Dienstleister für Hochschulen, außeruniversitäre Einrichtungen und die dort tätigen Wissenschaftler*innen, Publikationen auf Open-Access-Repositorien frei zugänglich zu machen und fördert das Zusammenspiel von wissenschaftlichen Einrichtungen und Verlagen. DeepGreen wurde von Januar 2016 bis Juni 2021 von der Deutschen Forschungsgemeinschaft gefördert und wird nun vom Kooperativen Bibliotheksverbund Berlin-Brandenburg, von der Bayerischen Staatsbibliothek und von der Universitätsbibliothek Erlangen-Nürnberg in arbeitsteiliger Eigenleistung für zwei Jahre weiterbetrieben. Der vorliegende Beitrag beleuchtet vielfältige Aspekte bei der Realisierung von DeepGreen und geht auf die Perspektiven dieser zentralen Open-Access-Infrastruktur für deutsche Wissenschaftseinrichtungen ein.

Summary

DeepGreen is an infrastructure that collects journal articles from academic publishers and sends them to authorised libraries for publication in their repositories. As a service provider for universities, non-university institutions and the academics working there, DeepGreen supports libraries in making publications freely accessible on Open Access (OA) repositories and promotes the interaction between scientific institutions and publishers. DeepGreen was funded by the German Research Foundation from January 2016 to June 2021 and will be operated for two more years by the Cooperative Library Network Berlin-Brandenburg, the Bavarian State Library and the Erlangen-Nuremberg University Library. The article discusses various aspects of the realisation of DeepGreen and addresses the prospects of this central OA infrastructure for German academic institutions.

Zitierfähiger Link (DOI): <https://doi.org/10.5282/o-bib/5764>

Autorendentifikation:

Boltze, Julia: ORCID: <https://orcid.org/0000-0002-0819-4271>;

Höllerl, Annika: GND: [1252420587](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9);

Kuberek, Monika: GND: [1186002352](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63862-p0011-9); ORCID: <https://orcid.org/0000-0002-1672-5271>;

Lohrum, Stefan: GND: [1252421672](#);

Pampel, Heinz: GND: [1088453392](#); ORCID: <https://orcid.org/0000-0003-3334-2771>;

Putnings, Markus: GND: [1043244379](#); ORCID: <https://orcid.org/0000-0002-6014-9048>;

Rusch, Beate: ORCID: <https://orcid.org/0000-0001-7664-4097>;

Schäffler, Hildegard: GND: [118181408](#);

Söllner, Konstanze: GND: [1058102818](#); ORCID: <https://orcid.org/0000-0002-6263-7846>

Schlagwörter: Open Access; Open-Access-Lieferdienst; Zweitveröffentlichung

Dieses Werk steht unter der [Lizenz Creative Commons Namensnennung 4.0 International](#).

1. Einleitung

Die Ursprungsidee von DeepGreen bestand darin, bereits vorhandene grüne Open-Access-Rechte in den von der Deutschen Forschungsgemeinschaft (DFG) geförderten Allianz-Lizenzen systematisch umzusetzen und die entsprechenden Veröffentlichungen möglichst automatisiert an die berechtigten Repositorien zu vermitteln. Während die Allianz-Lizenzen vielfach zu Open-Access-Transformationsverträgen weiterentwickelt wurden, sehen sich viele Einrichtungen weiterhin mit der Herausforderung der automatisierten Speicherung von Publikationen auf Open-Access-Repositorien konfrontiert. Mit DeepGreen wurde eine Infrastruktur geschaffen, die Bibliotheken als Dienstleister für Hochschulen, außeruniversitäre Einrichtungen und die dort tätigen Wissenschaftler*innen dabei unterstützt, Publikationen auf Open-Access-Repositorien frei zugänglich zu machen.

Aufgebaut und etabliert wurde eine Infrastruktur, die Zeitschriftenartikel (d.h. Datenpakete bestehend aus dem Volltext und den zugehörigen Metadaten) von wissenschaftlichen Verlagen abholt und berechtigten Bibliotheken zur Veröffentlichung in ihren Repositorien sendet. Damit unterstützt DeepGreen das Zusammenspiel von wissenschaftlichen Einrichtungen und Verlagen. Das Vorhaben wurde von Januar 2016 bis Juni 2021 von der DFG als Verbundprojekt gefördert. Beteiligt waren der Kooperative Bibliotheksverbund Berlin-Brandenburg (KOBV), die Bayerische Staatsbibliothek (BSB) mit der Verbundzentrale des Bibliotheksverbands Bayern (BVB), das Helmholtz Open Science Office sowie die Universitätsbibliotheken der Technischen Universität Berlin (TUB) und der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU).

In der ersten Förderphase (2016–2017) wurden erfolgreich ein technischer Prototyp, basierend auf der Open-Source-Software des Publications Routers¹ des Joint Information Systems Committee (Jisc)², entwickelt und erste Verlage als Partner gewonnen. Im Mittelpunkt der zweiten Förderphase stand ein produktionsvorbereitender Testbetrieb von DeepGreen, in dessen Verlauf neun wissenschaftliche Verlage, 63 institutionelle Repositorien und drei Fachrepositorien eine Kooperation mit DeepGreen eingingen. Die Rückmeldungen der Testpartner*innen führten zu zahlreichen Funktionserweiterungen, einer stärkeren Profilierung des Verlagsportfolios und nicht zuletzt zu Empfehlungen

1 Jisc Publications Router: <https://pubrouter.jisc.ac.uk/>, Stand: 15.09.2021.

2 Joined Information Systems Committee, <https://www.jisc.ac.uk/>, Stand: 15.09.2021.

für die optimierte Übernahme und Weiterverarbeitung in den Repositorien. Im Testbetrieb wurden zwischen September 2019 und Juni 2021 35.083 Publikationen erfolgreich zugestellt.

Im Juni 2021 ist DeepGreen in einen zweijährigen Pilotbetrieb übergegangen, der von KOBV, BSB und FAU getragen und aus Eigenmitteln finanziert wird. Ab Mitte 2023 soll ein Betriebskonzept greifen, das eine finanzielle Beteiligung durch die teilnehmenden Bibliotheken vorsieht. Die Zielgruppe von DeepGreen stellen hierbei alle überwiegend öffentlich finanzierten Hochschulen und Universitäten mit institutionellen Repositorien in Deutschland, sowie außeruniversitäre Forschungseinrichtungen und Fachrepositorien dar.

Der vorliegende Beitrag beleuchtet vielfältige Aspekte bei der Realisierung von DeepGreen und geht auf die Perspektiven dieser zentralen Open-Access-Infrastruktur für deutsche Wissenschaftseinrichtungen ein.

2. Technik

Technisch funktioniert DeepGreen wie eine Datendrehscheibe: Verlage liefern Volltexte und Metadaten im NISO JATS Format³ über sFTP und DeepGreen verarbeitet diese Daten. Die Daten durchlaufen verschiedene Prüfroutinen und DeepGreen teilt sie schließlich berechtigten Institutionen zu. Jeder Artikel im Konto einer Institution erhält als eindeutige Zuweisung eine sogenannte Notifikation. Die Institutionen haben die Möglichkeit, die Artikel via OAI-PMH⁴ und REST-API⁵ abzuholen oder sie über SWORD⁶ an ihr Repository liefern zu lassen. Da die verschiedenen, von Repositorien eingesetzten Softwaretypen für den Datenimport unterschiedliche Metadatenformate benötigen, kann DeepGreen auf Wunsch die JATS-Daten in METSMODS oder OPUS 4-XML formatieren (siehe Abbildung 1).

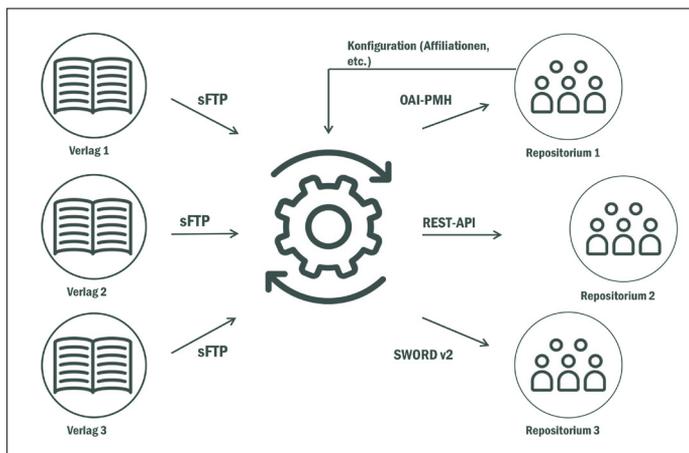


Abbildung 1: Funktionsweise der DeepGreen-Datendrehscheibe

3 NISO JATS: <<https://www.niso.org/standards-committees/jats>>, Stand: 15.09.2021.

4 OAI-PMH: <<https://www.openarchives.org/pmh/>>, Stand: 15.09.2021.

5 Representational State Transfer: <https://de.wikipedia.org/wiki/Representational_State_Transfer>, Stand: 15.09.2021.

6 SWORD: <[https://en.wikipedia.org/wiki/SWORD_\(protocol\)](https://en.wikipedia.org/wiki/SWORD_(protocol))>, Stand 15.09.2021.

2.1 Verlage und Datendrehscheibe

Eine der Grundfunktionen von DeepGreen besteht in der eindeutigen Zuordnung der Artikel sowohl zu einer Lizenz als auch zu einer Institution. Die Institutionszugehörigkeit ergibt sich aus der Angabe zur Affiliation der Autor*innen. Diese ist zwar in den Metadaten vorhanden, allerdings in der Regel nicht in standardisierter Form. Dieses Problem wird mit Hilfe einer sogenannten Affiliationsdatei gelöst. Teilnehmende Institutionen bekommen ein eigenes Konto und hinterlegen in dieser Datei möglichst viele unterschiedliche Namensvariationen ihrer eigenen Institution. Diese Namensvariationen werden mit den Affiliationsangaben in den Artikel-Metadaten verglichen. Bei einem Match wird der Artikel unter der Voraussetzung, dass die Institution berechtigt ist, diesen Artikel zu bekommen, zugestellt. Perspektivisch soll die Identifikation der Institutionen durch institutionelle Identifier wie ROR (Research Organization Registry⁷) erfolgen. Die Verwendung eines solchen persistenten Identifikators würde die Zuordnung deutlich vereinfachen, doch leider werden diese in den Verlagsmetadaten noch nicht flächendeckend verwendet.

Für die lizenzrechtliche Prüfung insbesondere beim grünen Open Access müssen weitere Datenquellen hinzugezogen werden, da die Metadaten die benötigten Informationen nicht enthalten. Die Prüfung zur Berechtigung einer Institution in Folge der Teilnahme an einem entsprechenden Lizenzvertrag erfolgt auf Datenbasis der Elektronischen Zeitschriftenbibliothek (EZB)⁸. Bevor ein Verlag Artikel liefert, wird in DeepGreen eine Liste mit ISSNs dieses Verlags hinterlegt (Kollektionsliste). Zusätzlich enthält die Kollektionsliste den einschlägigen Lizenzzeitraum pro Zeitschrift (bspw. der Allianz- oder Nationallizenzen). Des Weiteren wird eine Liste der lizenznehmenden Bibliotheken (Teilnehmendenliste) mit der Zeitschriftenliste verknüpft. Enthält die Kollektionsliste ausschließlich Titel, die Gold Open Access sind, wird die Teilnehmendenliste nicht benötigt.

Diese Zuordnungslogik wurde während der Projektlaufzeit auf Basis von Anforderungen der teilnehmenden Einrichtungen immer wieder um Funktionen erweitert. So ist DeepGreen seit Ende 2020 in der Lage, Artikel eines definierten Fachgebietes an disziplinäre Repositorien zu senden. Die fachliche Zuordnung basiert auf einer Auswahl von Zeitschriftentiteln, die das jeweilige Fachrepositorium definiert und die dann als Titelliste in DeepGreen hinterlegt wird. Aktuell nimmt das Interesse von Fachrepositorien speziell aus dem Kontext der DFG-geförderten Fachinformationsdienste für die Wissenschaft (FID) zu. Sollte es auf Verlagsseite künftig zu weiteren Standardisierungsaktivitäten kommen, beispielsweise durch den Einsatz einer international anerkannten Klassifikation auf Artikelebene, könnte eine fachliche Zuordnung auch auf dieser Datenbasis durchgeführt werden.

Seit August 2021 haben die an DeepGreen beteiligten Bibliotheken in ihrem Konto die Möglichkeit, individuell zu entscheiden, von welchen Verlagen sie Artikel bekommen möchten und von welchen nicht. Durch diese Filterfunktion wird gleichzeitig transparenter, welche Lizenzinformationen in DeepGreen hinterlegt sind, da die Institutionen so einsehen können, für welche Lizenzen und Verlage sie Artikel beziehen können (siehe Abbildung 2).

7 Research Organization Registry: <<https://ror.org/>>, Stand: 15.09.2021.

8 Elektronische Zeitschriftendatenbank: <<https://ezb.uni-regensburg.de/zeit/>>, Stand: 15.09.2021.

Filter

Your institution appears in the following licenses. [More information here](#)

Please uncheck all licenses you don't want to receive articles from:

- Sage Journals Komplett AL [EZB-NALIW-00495]
- FID Pharmazie Future Science [FID-FUTURES-00001]
- Karger Journals Komplett NL [EZB-NALJB-00504]
- SAGE Journals NL [EZB-NALJC-00505]
- Sage Journals Humanities and Social Science Upgrade 2015-16 [EZB-NALIW-01314]
- Wiley Open Access Journals [OA-WILEY-00001]
- Wiley DEAL KP [EZB-WIDEA-01707]
- Frontiers Open Access Journals [PUB-FRONT-00001]
- Gold: Hogrefe [PUB-HOGRE-00001]
- MDPI Open Access Journals [PUB-MDPI0-00001]
- Gold: Karger OA Journals [PUB-KARGER-00001]

Abbildung 2: Ansicht der Filterfunktion in den DeepGreen-Konten der Repositorien

Insgesamt ist der Bezug der Lizenzinformationen für DeepGreen noch nicht zufriedenstellend gelöst. So speichert die EZB keine historischen Lizenzinformationen, die für die Verarbeitung retrospektiver Verlagslieferungen an DeepGreen notwendig wären. Zudem wird durch die EZB nicht erfasst, dass bei Austritt einer Institution aus Lizenzverträgen nach wie vor Open-Access-Rechte an älteren Zeitschriftenjahrgängen bestehen. DeepGreen beschränkt sich in der Folge darauf, aktuelle Lizenzinformationen auszuwerten. Auch das Electronic Resource Management-System LAS:eR⁹ stellt hier derzeit keine praktikable Alternative dar, da bisher nicht absehbar ist, dass retrospektive Informationen aus Allianz- und Nationallizenzen in LAS:eR übertragen werden.

2.2 Datendrehscheibe und Repositorien

Die in den teilnehmenden Institutionen eingesetzten Repositorien setzen auf diversen Software-Systemen auf und bringen unterschiedliche Voraussetzungen für die Übernahme und Verarbeitung der von DeepGreen gelieferten Artikel mit. In Zusammenarbeit mit den Institutionen wurden im Laufe des Projekts die lokalen Bedingungen evaluiert, die Schnittstellen der Datendrehscheibe angepasst und erweitert und eine Handreichung¹⁰ mit Implementierungsanleitungen für Schnittstellen und Workflows auf Repositorienseite erstellt. Repositorien, die mit den gängigen Software-Lösungen DSpace, EPrints, MyCoRe oder OPUS 4 arbeiten, können an die Datendrehscheibe angeschlossen werden.¹¹

Der Workflow zur Verarbeitung einer Notifikation auf Repositorienseite unterscheidet sich – je nach Art des Repositoriums und den spezifischen Anforderungen der teilnehmenden Institution – insbesondere

9 LAS:eR, <https://www.hbz-nrw.de/produkte/digitale-inhalte/las-er>, Stand: 24.11.2021.

10 Bertelmann, Roland u.a.: DeepGreen. Open-Access-Transformation in der Informationsinfrastruktur. Anforderungen und Empfehlungen. Version 1.0, 2021. Online: <<https://doi.org/10.12752/8150>>, Stand: 15.09.2021.

11 Ebd., S. 14 ff.

im Grad der automatischen bzw. manuellen Verarbeitung, wobei allgemeine Arbeitsschritte identifiziert werden konnten.¹² Daraus wurden im Projekt Empfehlungen¹³ für die Institutionen und ein Workflow-Schema entwickelt, das universell – sowohl bei einer automatischen als auch bei einer manuellen Verarbeitung der Notifikationen und ebenso in den Fachrepositorien – angewandt werden kann (siehe Abbildung 3).

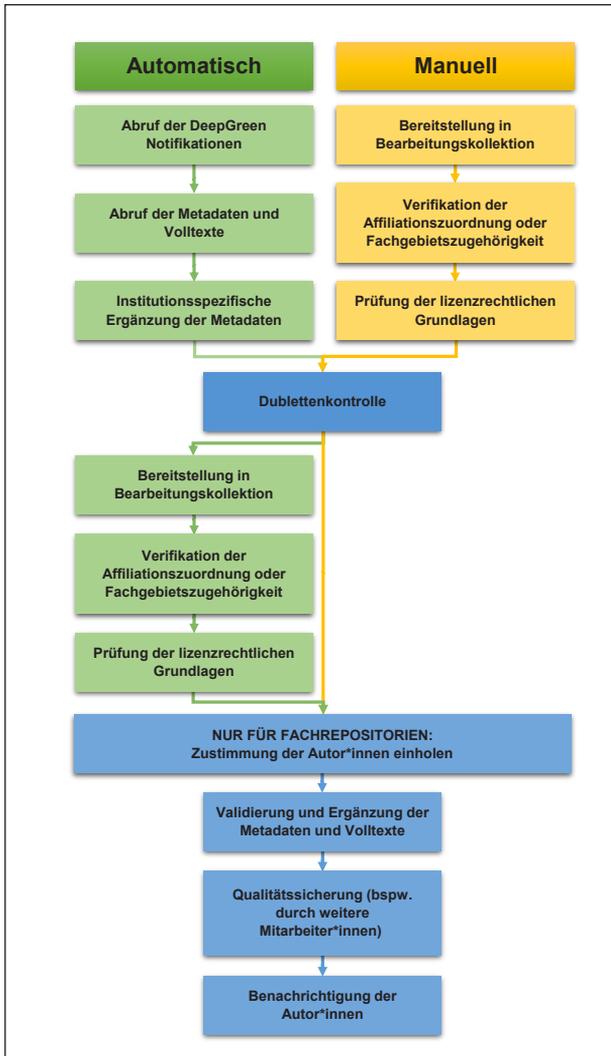


Abbildung 3: Schema der empfohlenen Workflows für die Verarbeitung von DeepGreen-Notifikationen durch Repositorien

12 Ebd., S. 19 ff.

13 Ebd., S. 22 ff.

Ein wesentlicher Arbeitsschritt bei der Verarbeitung der DeepGreen-Zustellungen ist neben der Prüfung der lizenzrechtlichen Grundlagen, für die die Bibliothek dem Verlag gegenüber haftet, die Erkennung von Dubletten und anderer Dokumentversionen. Diese Dublettenerkennung muss auf Repositorienseite erfolgen, da DeepGreen aufgrund institutionsspezifischer Anforderungen an den Umgang mit Dubletten keinen generischen Algorithmus zur Verfügung stellen kann. Eine eingehende Analyse der lokalen Möglichkeiten erbrachte, dass die in den Repositorien integrierten Suchmaschinen Apache Solr (Apache Software Foundation)¹⁴ und Elasticsearch (Elasticsearch B.V.)¹⁵ automatisierbare Funktionen für eine individualisierte Dublettenerkennung auf Basis von Prüfsummen anbieten. Im Ergebnis wird teilnehmenden Einrichtungen daher die Implementierung einer eigenen Dublettenerkennung empfohlen.¹⁶

2.3 Datendrehscheibe und Forschungsinformationssysteme

Die Belieferung von Forschungsinformationssystemen (FIS) konnte ohne größere zusätzliche technische Erweiterungen realisiert werden. Die mit den Forschungsinformationssystemen korrespondierenden Repositorien der folgenden Einrichtungen wurden erfolgreich an DeepGreen angebunden:

- Universität Bamberg
- Universität Bielefeld
- HAW Hamburg
- Technische Universität Hamburg
- Universität der Bundeswehr Hamburg
- Leibniz Universität Hannover

Die gewonnenen Erfahrungen bei der Anbindung wurden in der Handreichung dokumentiert.¹⁷

2.4 Weiterentwicklung der Datendrehscheibe

Seit 2021 wird DeepGreen durch den externen Dienstleister Cottage Labs¹⁸ weiterentwickelt. Mit dieser Entscheidung wurden die technisch administrative Betreuung der Datendrehscheibe und die funktionale Weiterentwicklung der Software in verschiedene Hände gelegt. Die Wahl fiel auf diese Firma, da sie die Basissoftware, den bereits erwähnten Jisc Publications Router, entwickelt hatte und dadurch bereits zu Beginn der Zusammenarbeit über umfassende Expertise verfügte. Mit dieser externen Unterstützung in Verbindung mit der Open-Source-Veröffentlichung der Software¹⁹ ist eine solide Basis für die nachhaltige Weiterentwicklung der Software nach Ende der Förderzeit gegeben.

2.5 Abgrenzung gegenüber anderen Datenlieferanten

Im Vergleich zur Zeit der ersten Förderphase (2016-2017) gibt es zwischenzeitlich große technische Fortentwicklungen sowohl auf Repositorienseite als auch auf Verlagsseite. Konkret können manche

14 Apache Solr: <<https://solr.apache.org/>>, Stand: 15.09.2021.

15 Elasticsearch: <<https://www.elastic.co/de/elasticsearch/>>, Stand: 15.09.2021.

16 Grundlegende Workflows und Konfigurationen in den Suchmaschinen Apache Solr, Elasticsearch und Elasticsearch in Kombination mit Logstash sind in der Handreichung beschrieben (siehe Bertelmann, Roland u.a., S. 28 ff.).

17 Bertelmann, Roland u.a.: DeepGreen. Open-Access-Transformation in der Informationsinfrastruktur. Anforderungen und Empfehlungen. Version 1.0, 2021. Online: <<https://doi.org/10.12752/8150>>, Stand: 15.09.2021.

18 Cottage Labs: <<https://cottagelabs.com/>>, Stand: 15.09.2021.

19 GitHub Repositoryum DeepGreen: <<https://github.com/oa-deepgreen>>, Stand: 15.09.2021.

Repositoriensoftwaretypen Schnittstellen zu anderen Datenlieferanten wie PubMed, Scopus, Web of Science, etc., nutzen und bibliographische Metadatenätze importieren. Umgekehrt bieten einige Verlage selbst auch SWORD-Ablieferungen an institutionelle Repositorien an. Die Vorteile von DeepGreen liegen in der Mitlieferung der Volltexte, der höheren Datenqualität, der Transformation (z.B. in OPUS 4-XML) zur Sicherstellung der Kompatibilität mit der jeweiligen Repositoriensoftware²⁰, der Rechtssicherheit, v.a. bei spezielleren Lizenzkontexten, und der schnelleren Datenlieferung im Vergleich zu den Indexverfahren von PubMed, Scopus und Web of Science.

3. Verlage

Neben den genannten technischen Entwicklungen kommt der Zusammenarbeit mit Verlagen in diesem Projekt eine wesentliche Rolle zu. Von Beginn an hat DeepGreen auf der technischen Seite unter der Maxime gearbeitet, den Aufwand für Verlage so gering wie möglich zu halten und deren Anforderungen weitgehend umzusetzen. Dass manche Verlage bereits mit Jisc kooperierten, war aufgrund der dadurch gegebenen ähnlichen technischen Bedingungen ein großer Vorteil.

Über die technischen Aspekte hinaus kann DeepGreen gegenüber den Verlagen auf weitere Vorteile verweisen: die bessere Verfügbarkeit und Verbreitung der Artikel, die leichtere Umsetzung von Verpflichtungen aus Verträgen oder Fördervorgaben und die Beschränkung von verlagsseitigen Datenlieferungen auf einen einzigen, zentralen Abnehmer. Auch hat sich DeepGreen in seiner ersten Projektphase ausschließlich auf Inhalte aus Allianz- und Nationallizenzen konzentriert. Die Open-Access-Klausel der Allianz-Lizenzen sieht Zweitveröffentlichungsrechte für Autor*innen und deren Institutionen explizit vor.²¹ DeepGreen begann so als ein Service, der bereits zugestandene Rechte in einer effektiveren und effizienteren Form als zuvor wahrzunehmen half.

In der Umsetzung dieses Vorhabens stieß DeepGreen teils auf die erwarteten, teils auf überraschende Hürden. Zunächst stellte sich bei einer Analyse der bestehenden Allianz-Lizenz-Verträge heraus, dass nicht alle Verträge tatsächlich mit der vollen Zweitveröffentlichungs-Klausel, die auf die publizierte Version abhebt, ausgestattet waren, sondern stärker eingeschränkte Rechte beinhalteten. Dadurch verringerte sich die Zahl der in Frage kommenden Verlage entsprechend. In der Gruppe der angesprochenen Verlage stieß man zusätzlich auf Hürden. Mangelnde technische und personelle Kapazitäten für die Umsetzung der Datenlieferungen waren der häufigste Grund einer Nicht-Teilnahme an DeepGreen. Vielfach gestalteten sich auch bei vorhandenem Interesse die Gespräche wegen der notwendigen Rückläufe unter Einschluss der Rechtsabteilungen durchaus langwierig. Mit drei Häusern konnte schließlich in der ersten Phase eine erfolgreiche Zusammenarbeit aufgebaut werden: Karger, SAGE und De Gruyter, die auch weiterhin Kooperationspartner sind.

20 Dies können Verlage mit SWORD-Ablieferungen i.d.R. nicht leisten.

21 „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 disziplinspezifische Repositorien ihrer Wahl einzupflegen und im Open Access zugänglich zu machen. Das gleiche Recht besitzen die autorisierten Einrichtungen, denen die jeweiligen Autoren angehören.“ Deutsche Forschungsgemeinschaft: Grundsätze für den Erwerb DFG-geförderter überregionaler Lizenzen (Allianz-Lizenzen), 2015, S. 9. Online: <http://www.dfg.de/formulare/12_181/12_181_de.pdf>, Stand: 15.09.2021.

Im Betrieb wurde deutlich, dass die Allianz-Lizenz-Klausel nicht alle rechtlich relevanten Fragen im Kontext von DeepGreen hinreichend löst. So blieb zusätzlicher Regelungsbedarf bestehen – etwa hinsichtlich des Lizenz-Status der mitgelieferten Metadaten oder des Übergangs der Open-Access-Rechte aus der Allianz-Lizenz auf die Teilnehmer der Nationallizenz über die Moving Wall.²² Auch der ausschlaggebende Zeitpunkt zur Berechnung des Embargos war oft nicht definiert. Derartige Fragen konnten schließlich über die Kooperationsabkommen zwischen DeepGreen und den Verlagen geklärt werden.

Andere Verlage waren zwar durchaus interessiert, jedoch scheiterte die Kooperation an den technischen Voraussetzungen. Schon in einer Umfrage²³ unter zahlreichen Verlagen zum Projektstart von DeepGreen konnte erhoben werden, dass die in Verlagen vorliegenden Metadatenformate divers und einheitliche Standards nicht zu erwarten sind. DeepGreen verarbeitet Metadaten im NISO JATS-Format, das zwar eines der gängigsten Formate ist, jedoch hauptsächlich von großen Anbietern verwendet wird. Kleinere Häuser mit anderen Metadaten schemata waren somit für DeepGreen zunächst nicht erreichbar. Selbst bei der Verwendung von NISO JATS ergaben sich Schwierigkeiten z.B. durch die Verwendung verschiedener JATS-Dialekte. Auch bei den laufenden Kooperationen ergaben sich immer wieder Überraschungen in Bezug auf die Metadatenqualität, beispielsweise ungefüllte Tags zu Publikationsdatum und Verlag, die Anpassungen entweder auf Seiten von DeepGreen oder des Verlags notwendig machten.

Am meisten veränderte allerdings der Wandel der Publikationswege die Verlagsverhandlungen. In der Vorbereitung des Projektantrags 2015 war die Open-Access-Welt noch stärker auf grünen Open Access fokussiert. Mit den Allianz-Lizenzen hatte DeepGreen einen geeigneten Hebel, um auf Verlage zuzugehen und baute seinen Service in der Konsequenz als – wie der Name schon sagt – Angebot im Bereich der Zweitveröffentlichungsrechte, also des grünen Open Access auf. Der Publikationsmarkt jedoch entwickelte sich über die Laufzeit der ersten Projektphase hin weiter in Richtung goldenen und hybriden Open Access.

Auch wissenschaftspolitisch gewann die Open-Access-Transformation erheblich an Bedeutung: Im Jahr 2016 starteten die Verhandlungen im Projekt DEAL, in denen unter Federführung der Hochschulrektorenkonferenz mit den größten Wissenschaftsverlagen deutschlandweite Abschlüsse zu Publish & Read-Modellen angestrebt wurden. 2019 bzw. 2020 gelang dies erfolgreich mit den Verlagen Wiley und Springer Nature. Auch die DFG erweiterte zu diesem Zeitpunkt ihre Förderlinien: Neben die bislang DFG-geförderten Allianz-Lizenzen traten von 2017 bis 2020 Open-Access-Transformationsverträge mit einer eigenen Ausschreibung in den Fokus.

In Folge dieser Veränderungen entwickelte sich auch DeepGreen weiter. Das Angebot reiner Gold-Open-Access-Verlage, Inhalte aus Transformationsverträgen, wie auch die Lizenzen der

22 Die Inhalte der für die Teilnehmer kostenpflichtigen Allianz-Lizenz gehen Großteils nach Ablauf einer gewissen Zeitspanne, „Moving Wall“ genannt, in eine Nationallizenz über, deren Teilnahme für zugelassene Institutionen kostenfrei ist.

23 Becker, Pascal-Nicolas u.a.: Questionnaire for effective exchange of bibliographic metadata. Current status of publishing houses, 2016. Online: <[urn:nbn:de:0297-zib-60419](https://nbn-resolving.org/urn:nbn:de:0297-zib-60419)>, Stand: 15.09.2021.

Fachinformationsdienste für die Wissenschaft (FID) sowie grüner Open Access außerhalb der Allianz-Lizenzen erschien nun als denkbare Grundlage für den Verteilmechanismus von DeepGreen. Dadurch vergrößerte sich das Spektrum an möglichen Kooperationspartnern, aber auch die Komplexität der internen Prüfroutinen im DeepGreen-Router, die auf die unterschiedlichen Anforderungen abgestimmt werden mussten. Mit der Öffnung des Angebots konnten weitere Verlage ins Boot geholt werden: die beiden reinen Gold-Open-Access-Verlage Frontiers und MDPI, Wiley mit dem DEAL-Abschluss, Future Science als Beispiel einer FID-Lizenz sowie schließlich noch zum Ende der Projektlaufzeit Hogrefe, ebenfalls mit einem Transformationsvertrag. Verhandlungen mit weiteren Verlagen sind weit fortgeschritten.

Als wichtiges Kriterium zur Auswahl der Verlage hat sich das Publikationsaufkommen deutscher Wissenschaftler*innen in dem jeweiligen Verlag herausgestellt. Die Einrichtung einer automatisierten Lieferung über DeepGreen lohnt sich nur, wenn im Gegenzug mehrere Repositorien von regelmäßigen Lieferungen profitieren. Für vereinzelte Artikel, die nur für sehr wenige Institutionen in Deutschland Relevanz haben, ist der technische und organisatorische Aufwand hingegen kaum zu rechtfertigen. Als Beispiel hierfür sind die hochspezialisierten FID-Lizenzen zu nennen. Mit Future Science wurde zwar ein Verlag gefunden, mit dem automatisierte Lieferungen vereinbart werden konnten, die reine Zahl der erfolgten Zuteilungen spricht allerdings nicht dafür, das Modell auf andere FID-Lizenzen auszuweiten. Die gleiche Herausforderung stellt sich bei mittelständischen Verlagen, von denen zwar einige Interesse an DeepGreen angemeldet haben, bei denen die Größe des Zeitschriftenportfolios jedoch einen ebenfalls nur sehr überschaubaren Datenfluss verspricht.

Zusammenfassend lässt sich feststellen, dass sich trotz anfänglicher Zurückhaltung im Projekt DeepGreen stabile Verlagsbeziehungen entwickelt haben und weiteres Potential besteht. Als erfolgreich hat sich einerseits die Strategie herausgestellt, mit DeepGreen an bestehende Lizenzvereinbarungen anzuknüpfen, etwa im Rahmen der Allianz-, National- und FID-Lizenzen und deren Weiterentwicklung als Transformationsverträge, andererseits aber auch die Praxis, auf Gold-Open-Access-Verlage zuzugehen. Dabei hat sich gezeigt, dass der Automatisierungsprozess, den DeepGreen liefert, vor allem für große Verlage mit umfangreichem Zeitschriftenportfolio und standardisierten Metadaten eine geeignete Lösung darstellt. Perspektivisch könnte schließlich die Maschinenlesbarkeit von Lizenzinformationen aus den komplexer werdenden Verträgen, beispielsweise durch Nutzung von ONIX-PL²⁴, im Dialog mit Verlagen an Bedeutung gewinnen.

4. Verstetigung

Nicht zuletzt durch den ausgedehnten produktionsvorbereitenden Testbetrieb mit vielen Beteiligten, begleitet von einer intensiven Vortragstätigkeit in der Fachöffentlichkeit, ist es DeepGreen gelungen, in der deutschen Open-Access-Landschaft ein bekannter Name zu werden, wobei nicht vergessen werden darf, dass ein dauerhaftes, nachhaltig finanziertes Betriebsmodell noch im Detail etabliert werden muss.

24 ONIX-PL Working Group: <<http://www.niso.org/standards-committees/onix-pl-working-group>>, Stand: 15.09.2021.

4.1 Rechtsfragen

Die rechtlichen Themen, die das Verhältnis zwischen Verlagen, DeepGreen-Betreibern und Bibliothek betreffen, wurden bereits während der erweiterten Testphase behandelt, indem zwischen allen Beteiligten Verträge bzw. Haftungsausschlüsse abgeschlossen wurden. Im Dialog mit den Einrichtungen hat sich gezeigt, dass sich die Bibliotheken (und auch die Wissenschaftler*innen) DeepGreen als eine rechtlich möglichst verbindliche Clearingstelle wünschen, die idealerweise das alleinige Haftungsrisiko trägt. Dagegen steht, dass die Datenlieferungen auf Verträgen zwischen Verlagen und Bibliotheken basieren; die DeepGreen-Verträge setzen auf bestehenden Vereinbarungen auf und ersetzen diese nicht, so dass die Bibliotheken weiterhin haftender Vertragspartner den Verlagen gegenüber bleiben. Auch so mancher Verlag wünscht sich von DeepGreen eine stärkere Kontrollfunktion den Bibliotheken gegenüber, so dass sich hier der Dienst in einer klassischen Sandwichposition mit Erwartungen von unterschiedlichen Seiten wiederfindet.

Soll der Betrieb des Dienstes, wie in der Projektzeit erfolgreich eingeübt, weiterhin kooperativ-verteilt von mehreren Institutionen erbracht werden, muss auch dafür ein tragfähiger Rahmen gefunden werden. Nicht zuletzt stellt sich die Frage einer transparenten Governance bzw. der Einbindung der Anwender*innen.

4.2 Personalausstattung

Ein erfolgreicher Betrieb von DeepGreen umfasst weit mehr als die Bereitstellung und Weiterentwicklung einer technischen Infrastruktur. Es bestehen hohe Erwartungen von Seiten aller Beteiligten an eine stetige, transparente Information über Ziele, geplante und realisierte Neuerungen. Gewünscht werden Ansprechpartner*innen für technische und organisatorische Fragen und Best Practices²⁵ im laufenden Betrieb als auch ein gewisses politisches Engagement für Open Access und Open Science.²⁶ Eine anspruchsvolle Daueraufgabe besteht auch im Gewinnen weiterer Verlage und Repositorien für DeepGreen und in der Pflege bestehender Kooperationen. Alle diese Aufgaben setzen auf personeller Ebene unterschiedliche Qualifikationen voraus: neben technischer und (lizenz-)rechtlicher Expertise ist die Kenntnis von Open-Access-Policies in den Hochschulen bzw. seitens der Förderer und den praktischen Anforderungen von Repositorienbetreibern wünschenswert. Diese Aufgaben werden im Pilotbetrieb in verteilten Rollen von mehreren Institutionen übernommen: So verantwortet die KOBV-Verbundzentrale den technisch-administrativen Bereich und steht als erster Ansprechpartner zur Verfügung, gefolgt von der Bayerischen Staatsbibliothek, die die Federführung für alle Verlagsverhandlungen innehat, und der Friedrich-Alexander-Universität Erlangen-Nürnberg als primärer Testpartner auf der Repositorienseite und gut vernetzter Multiplikator.

4.3 Finanzierung

DeepGreen steht mit Blick auf die Finanzierung des angestrebten Dauerbetriebs vor einer ähnlichen Fragestellung wie nicht wenige andere Infrastrukturprojekte auch. Nach Ende der Anschubfinanzierung

25 Dazu werden auf der DeepGreen-Webseite auch multimedial z.B. mehrere Tutorials in Form von Screencasts angeboten <<https://info.oa-deepgreen.de/hilfe/#Screencasts>>, Stand: 08.02.2022.

26 So wurde im Workshop „Kick-Off für die Zukunft von DeepGreen: bisherige Erfolge und neue Ziele“ am 02.03.2021 der Wunsch geäußert, dass sich DeepGreen noch stärker als bisher an der Open Access-Debatte z.B. an SCOSS und insbesondere im Kontext des grünen Wegs von Open Access engagieren könnte.

aus Drittmitteln und einer durch die Projektpartner getragenen Pilotphase muss angesichts der für Basisdienste dieser Art systematisch fehlenden Grundfinanzierung ein tragfähiges Geschäftsmodell gefunden werden, das den dauerhaften Betrieb ermöglicht.

Erfolgreiche Beispiele sind die Elektronische Zeitschriftendatenbank, die Beiträge von Teilnehmenden nimmt, nachdem der Dienst viele Jahre aus Projektmitteln und in Eigenleistung getragen wurde, oder auch der im Wesentlichen von Bibliotheken getragene Verein Kitodo, dessen Mitglieder die Grundkosten für die Weiterentwicklung der Software teilen und freiwillige Beiträge in einen Entwicklungsfonds einzahlen.

Für DeepGreen bietet sich ein Umlagemodell an, das beispielsweise eine Kostendifferenzierung nach Typ und Größe der Einrichtung im Sinne von Vollzeitäquivalenten vornimmt. Ob die Bibliotheken allerdings bereit sind, sich an einem Vollkostenmodell für DeepGreen zu beteiligen, ist zum gegenwärtigen Zeitpunkt offen.

5. Ausblick

DeepGreen birgt grundsätzlich das Potential, infrastrukturell an verschiedene etablierte und jüngere Services der Open-Access-Landschaft anzudocken oder diese zu unterstützen. Als Beispiel kann die Kooperation mit Wiley genannt werden, die aus dem Projekt DEAL hervorgegangen ist. Diese ließe sich mühelos auf andere Verlage bzw. Transformationsprojekte übertragen.

Mit dem Wandel hin zu Open Science wird ein breiter Blick auf die gesamte Forschungsleistung eingenommen. Über die textuelle Publikation, z.B. in Form eines Zeitschriftenaufsatzes hinaus, werden Forschungsdaten und Programmcode, die mit dem Artikel in Verbindung stehen, in den Blick genommen. Damit stellt sich auch die perspektivische Frage, wie sich DeepGreen mit seinem bisherigen Fokus auf Zeitschriftenartikel in eine zukünftige Open-Science-Infrastrukturlandschaft einordnet und wie die Dissemination und Vernetzung von digitalen Informationsobjekten technologisch möglich wäre.

Es erscheint zudem wahrscheinlich, dass neben der Kerndienstleistung, der Zustellung der Publikationen, künftig auch Monitoringfunktionen und damit verbundene Analysen rund um Open Access von Interesse werden können. Perspektivisch sind hier Zusammenspiele mit Diensten wie dem Open Access Monitor²⁷ oder auf internationaler Ebene Unpaywall²⁸ denkbar.

Zusätzlich ließe sich DeepGreen auch aktiv in Fördererentscheidungen einbinden: Beim in Deutschland noch relativ selten genutzten²⁹ Open Access Switchboard (OAS)³⁰ können neben Kosten- und Förderinformationen auch Wünsche der Institution bzw. des Förderers hinsichtlich einer Deponierung des Artikels in einem Repository hinterlegt werden (siehe Abbildung 4).

27 Open Access Monitor: <<https://open-access-monitor.de/#/home>>, Stand: 15.09.2021.

28 Unpaywall: <<https://unpaywall.org/>>, Stand: 15.09.2021.

29 Derzeit sind nur die Universitätsbibliothek Erlangen-Nürnberg und die Max Planck Digital Library (MPDL) OAS-Vertragspartner.

30 Open Access Switchboard: <<https://www.oaswitchboard.org/>>, Stand: 15.09.2021.

Article	Vor	Publication	open access / transformative journal
		License	CC BY-NC-SA
		Deposition	open repository, like PMC
		Researchdata	not applicable

Abbildung 4: Screenshot aus einer "OA Switchboard Eligibility Enquiry request"-E-Mail.

Es wäre denkbar, an dieser Stelle auch die Einspeisung in eine nationale Datendrehscheibe wie Jisc Publications Router oder DeepGreen zu kommunizieren. Die Universitätsbibliothek Erlangen-Nürnberg ist Mitglied im OAS Client Advisory Board und bringt dort die DeepGreen-Interessen ein.

Das Open Access Switchboard ist nur einer von mehreren intermediären Instanzen, die sich derzeit auf europäischer bzw. internationaler Ebene konstituieren, um Open Access Services und Daten für verschiedene Akteure (z.B. Autorinnen und Autoren, Verlage, Bibliotheken, Langzeitarchive) zu bündeln. DeepGreen erlaubt durch seine multivariate Ausrichtung im Lizenz- und Technikkontext eine Anbindung sowohl direkt an die Akteure als auch die Intermediäre.

DeepGreen hat somit das Potential, sich von dem ursprünglich geplanten Dienst, der die Umsetzung der grünen Open-Access-Komponente der Allianz-Lizenzen zum Ziel hatte, zu einem wesentlichen Baustein in der deutschen Open-Science-Landschaft zu entwickeln. DeepGreen versorgt im Rahmen des Test- und Pilotbetriebs eine große Zahl von Testpartnern mit hochwertigen Daten und kann im Dauerbetrieb dazu beitragen, dass sich ein umfassenderes deutsches Netzwerk im Sinne von Next Generation Repositories entwickelt. Es hängt nun von den deutschen Verbänden und den sie tragenden Bibliotheken ab, wie DeepGreen in deren Dienstleistungsportfolio integriert wird. Der Bedarf steigt mit jeder einzelnen Open-Access- und Transformationsvereinbarung.

Literaturverzeichnis

- Becker, Pascal-Nicolas u.a.: Questionnaire for effective exchange of bibliographic meta-data. Current status of publishing houses, 2016. Online: <[urn:nbn:de:0297-zib-60419](http://nbn-resolving.org/urn:nbn:de:0297-zib-60419)>, Stand: 15.09.2021.
- Bertelmann, Roland u.a.: DeepGreen. Open-Access-Transformation in der Informationsinfrastruktur. Anforderungen und Empfehlungen. Version 1.0, 2021. Online: <<https://doi.org/10.12752/8150>>, Stand: 15.09.2021.
- Deutsche Forschungsgemeinschaft: Grundsätze für den Erwerb DFG-geförderter überregionaler Lizenzen (Allianz-Lizenzen), 2015. Online: <http://www.dfg.de/formulare/12_181/12_181_de.pdf>, Stand: 15.09.2021.