DeepGreen – Pushing on the Green Road of Open Access

Eike Wannick, eike.wannick@os.helmholtz.de
Julia Boltze, boltze@zib.de

Agenda

1. The Green Road of Open Access in Germany

2. What is DeepGreen?

3. Technical Functionality: Requirements for Repositories

4. Summary and Outlook

1. The Green Road of Open Access in Germany

Where we start from…

- Development of a vivid Open Access publishing environment

- Increasing number of political Open Access guidelines

- A plural landscape of institutional and disciplinary repositories at libaries, universities and research institutions

- DFG (German Research Foundation) provides funding for research projects and supports the Open Access transformation

- Green Open Access component included in DFG-funded Alliance Licenses → unfortunatley this option often remains unused

- Due to this complex environment, a strong cooperation between all stakeholders (publishers, funding institutions, libraries, scientists…) necessary

The Open Access Component in Alliance Licenses

"Upon request, the licensor is obligated to physically supply the licensee with the complete product at no additional charge, i.e. including metadata and all full text, including digital objects that are part of the product, on suitable storage media and in suitable data formats as agreed.

Licensees may use the data provided to them in any way they deem suitable in order to make the product accessible to authorised users, in compliance with the license agreements. They may, for this purpose, integrate the data in technical usage/storage systems (hosting and archiving) operated by themselves or by third parties."

http://www.dfg.de/service/bildarchiv/index.html
https://www.dfg.de/formulare/12_181/12_181_en.pdf

Challenges

Time-consuming analysis of publications with regard to Alliance Licenses:

- High manual effort
- individual identification of authorized articles
- Less automization
- Error-prone

The green Open Access component in Alliance and National Licenses are often unused due to the high effort for libraries and research institutions.

Solution

Design a legally watertight, automated process:

- Assigning articles to the correct institutions (by using all affiliations in the metadata)
- Check whether the institution is allowed to get the specific article (Are there valid contracts?)
- Following embargo periods

Et voilà: DeepGreen

2. What is DeepGreen?

Lowering the Barriers for Green Open Access

The project DeepGreen

- Funded by the German Research Foundation (DFG)
- Consortium of six institutions
- First project phase: January 2016 – December 2017
- Second project phase: August 2018 – July 2020
- Cooperating publishers:
  - Initial cooperation partners: S. Karger, SAGE Publications
  - Later on: BMJ, De Gruyter, MDPI

- Screencasts explaining the functionality in some more detail are available at
  https://deepgreen.kobv.de/de/deepgreen/screencasts/
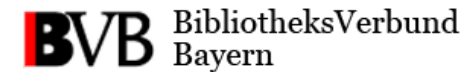
Aims of DeepGreen

- Lower the barriers for Green (secondary) Open Access publishing for repositories

- Create a legally watertight, highly automated process for delivering metadata and fulltext publications from publishers to repositories

- Target group: all mainly publicly funded libraries, research institutions and universities in Germany
  - About 250 with valid licenses with the publishing partners S. Karger und SAGE Publications
  - The more publishers participate, the more libraries and institutions can benefit from the service

What DeepGreen is NOT

DeepGreen (as a prototype) will not…

- be a repository or a dark archive
- enrich or modify the content metadata
- take legal responsibility
- Filter input data (e.g. duplicates)

→ DeepGreen is a push-forward system

# 3. Technical Functionality: Requirements for Repositories

Technical Requirements

- Ability to process a steady delivery of data (metadata & .pdf)

- Use of a flexible data model for license information

- Provide (all) requested interfaces for easy data exchange with publishers (in) and repositories (out)

- Milestones already achieved: OPUS4, DSpace and eSciDoc/Pubman are served by DeepGreen

Automated Workflow

Highly automated workflow with a central data router, which takes articles from publishers, figures out all the affiliations in the metadata and delivers them to entitled repositories.

- Based on an open source version of Jisc Publications Event Router (Open Source)
- Simple configuration
- Detailed reports
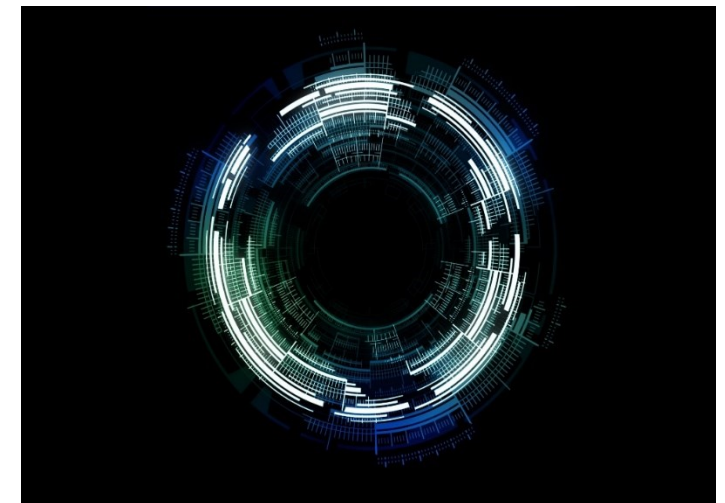
System Architecture of DeepGreen

Based on Jisc Publications Event Router
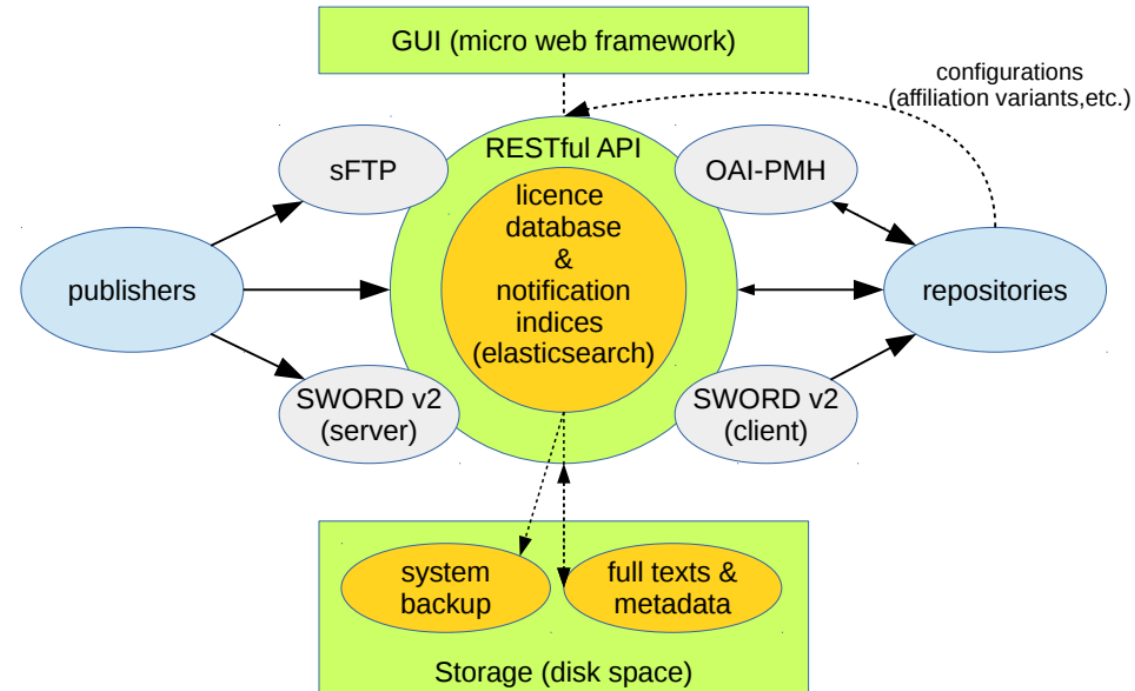
The fruitful cooperation with Jisc emphasizes the positive effect of Open Science.
https://github.com/JiscPER/jper/docs/system/ArchitectureOverview.png

Matching Process

Two-step procedure for each incoming article

1

Verification of licenses for the related journal via EZB-Database

- DeepGreen checks if a Alliance License is recorded for the journal (by using the ISSN of the journal, included in the metadata)
- Records of Alliance License collections of journal titles, volumes and corresponding qualified institutions are obtained from EZB database

2

Determination of entitled institutions by analysing the affiliation included in the metadata of the publication

- DeepGreen tries to match the affiliations found in article's metadata with affiliation snippets of all entitled institutions
- All matches and all non-matches are logged

DEEP GREEN
OPEN ACCESS TRANSFORMATION

Side note: What is the EZB database?

- German database for journal license information
- Contains information about journal holdings and licenses of participating libraries

https://rzblx1.uni-regensburg.de/ezeit/

The Affiliation Data

	A	B	C	D	E	F	G
1	Name Variants	Domains	Grant Numbers	Dummy1	Dummy2	Keywords	
2	Academia Fridericiana Erlangensis						
3	Academia Friderico Alexandrina Erlangen-Nürnberg						
4	Academia Friderico-Alexandrina						
5	Academia Regia Bavarica Friderico-Alexandrina						
6	Academia Regia Friderico-Alexandrina						
7	Bayerische Friedrich-Alexanders-Universität						
8	F.A.U. Erlangen-Nürnberg						
9	FAU Erlangen-Nürnberg						
10	Friedrich Alexander University						
11	Friedrich-Alexander-Universität Erlangen						
12	Friedrich-Alexander-Universität Erlangen-Nürnberg						
13	Friedrich-Alexander-Universität zu Erlangen						
14	Friedrich-Alexander-University Erlangen						
15	Friedrich-Alexanders-Universität						
16	Friedrichs-Akademie						
17	Königlich Bayerische Friedrich-Alexanders-Universität						
18	Rektor der Friedrich-Alexander-Universität Erlangen-Nürnberg						
19	Rektorat der Friedrich-Alexander-Universität Erlangen-Nürnberg						
20	Univ. Erlangen-Nürnberg						
21	Universidad de Erlangen-Núremberg						
22	Universidad de Erlangen-Nürnberg						
23	Universitas Literarum Regia Friderico-Alexandrina						
24	Università di Erlangen-Nürnberg						
25	Universität Erlangen						
26	Universität Erlangen-Nürnberg						
27	University Erlangen-Nuremberg						
28	University of Erlangen-Nuremberg						
29	University of Erlangen-Nürnberg						
30							
31							

Important fields:

Name variations
Domains
Grant numbers
Keywords

The Affiliation Dataset (UTF-8-encoding)

- **Name Variants,Domains,Grant Numbers,Dummy1,Dummy2,Keywords**
- ,fau.de,,,,
- ,uk-erlangen.de,,,,
- ,uni-erlangen.de,249169,,,
- Academia Fridericiana Erlangensis,,,,,
- Academia Friderico-Alexandrina,,,,,
- Academia Regia Bavarica Friderico-Alexandrina,,,,,
- Academia Regia Friderico-Alexandrina,,,,,
- Bayerische Friedrich-Alexanders-Universitä t,,,,,
- F.A.U. Erlangen-Nü rnberg,,,,,
- FAU Erlangen-Nü rnberg,,,,,
- Friedrich Alexander University,,,,,
- Friedrich-Alexander-Universitä t Erlangen,,,,,
- Friedrich-Alexander-Universitä t Erlangen-Nü rnberg,,,,,
- Friedrich-Alexander-Universitä t zu Erlangen,,,,,
- Friedrich-Alexanders-Universitä t,,,,,
- Friedrich-Alexanders-Universitä t zu Erlangen,,,,,
- Friedrichs-Akademie,,,,,
- Kö niglich Bayerische Friedrich-Alexanders-Universitä t,,,,,
- Universidad de Erlangen-Núremberg,,,,,
- Universitas Literarum Regia Friderico-Alexandrina,,,,,
- Universitä t Erlangen,,,,,
- Universitä t Erlangen-Nü rnberg,,,,,
- University of Erlangen-Nuremberg,,,,,

Summary from a repository perspective

- Research institutions / repositories get an DeepGreen account when…
    - a valid license exists (Alliance Licenses or Gold Open Access)
    - a EZB-ID of the institution exists

- A repository account in DeepGreen…
    - manages matching-criteria (affiliation file of the institution)
    - lists and provides the assigned publications/articles
    - allows configuration for a SWORD-interface

- Possible matching-criteria for a DeepGreen account are
    - Name variations of the institution, IP-Domains, grant numbers (as they appear for example in publications)
    - Institutions are responsible for managing the affiliation file and verifying assigned articles

3. Summary and Outlook

Start of the advanced test phase of DeepGreen

- In summer 2019 DeepGreen will start a testing phase with around 30 repositories.

- The focus is on the OA rights of the Alliance Licenses, but gold OA articles will be processed too.

- The aim of the testing phase is the identification of bugs and getting experience with data and repositories in a live simulation.

- Communication and agreements with more interested publishers are intended.

- S. Karger, SAGE Publications will participate and there is promising communication with MDPI, BMJ, Frontiers und Walter de Gruyter

The role of publishers

- The success of DeepGreen is depending on the willingness of publishers to participate

- For publishers DeepGreen offers the opportunity to be an active driver of the Green Road of Open Access and a pioneer within the Open Access transformation

- Interested publishers are always welcome to contact us!

Aims for the second funding phase

- Integrating disciplinary repositories and current research information systems (CRIS) into the DeepGreen workflow

- Widening the scope on other licensing models, as e.g. Gold Open Access

- Establish cooperations with more publishers

- Development of a business plan and mode of operation for a DeepGreen Service following the second funding phase

International perspective

- As DeepGreen aims to not only deliver Open Access Articles in the green road of Open Access and within the context of Alliance Licenses but also could work for OA Gold, a similar workflow could be interesting in an international context.

- Promising communication with Austria

- Pushes standarization and open data formats

Follow us to the Green Road!

Questions?
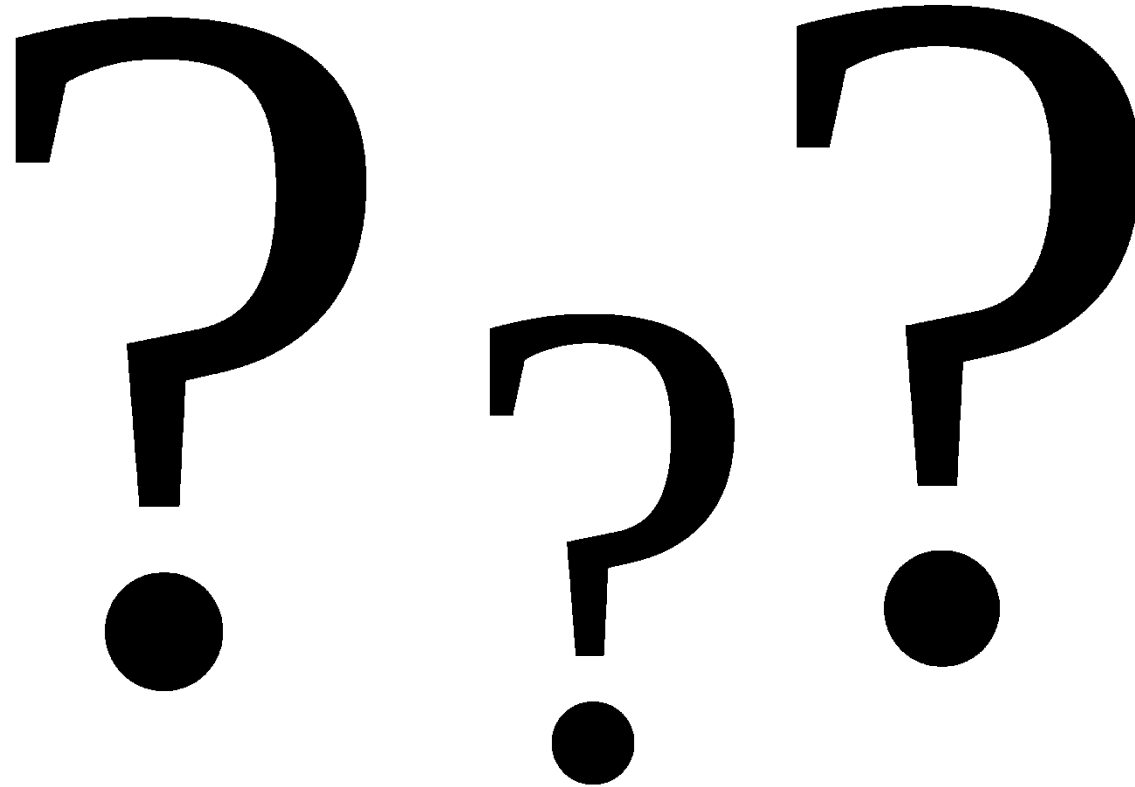
Learn more: Our publications

- **Christof, J., Bertelmann, R., Griebel, R., Koch, T. & Söllner, K.** (2016): DeepGreen - Entwicklung eines rechtssicheren Workflows zur effizienten Umsetzung der Open-Access-Komponente in den Allianz-Lizenzen für die Wissenschaft. ZIB-Report, 15-58. http://nbn-resolving.de/urn:nbn:de:0297-zib-56799.

- **Dierkes, T.** (2017): Ist die Umsetzung der Open-Access-Komponente in Allianz-Lizenzverträgen wirklich so kompliziert? ABI Technik , 37 (3), 237-237. https://doi.org/10.1515/abitech-2017-0051.

- **Goltz, J. A., Scheliga, K. S.** (2017): DeepGreen: Prototyping an efficient technical implementation of the open access components included in the Alliance licenses - Project presentations, Open Science Conference (Berlin 2017).

- **Koch, T., Ceynowa, K., Söllner, K., Christoph, J., Bertelmann, R.** (2018a): 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. http://nbn-resolving.de/urn:nbn:de:0297-zib-69612.

- **Koch, T., Rusch, B., Dierkes, T., Goltz-Fellgiebel, J. A., Schwidder, J., & Boltze, J.** (2018b). DeepGreen – Open Access Transformation. Zenodo. http://doi.org/10.5281/zenodo.1468934.

- **Putnings, M. & Rusch, B.** (2016): DeepGreen – Entwicklung eines rechtssicheren Workflows zur effizienten Umsetzung der Open-Access-Komponente in den Allianz-Lizenzen für die Wissenschaft. o-bib : das offene Bibliotheksjournal, 3 (4), 110 – 119. https://doi.org/10.5282/o-bib/2016H4S110-119.

- **Goltz-Fellgiebel, J.A., Putnings, M.** (2019): Open-Access-Transformation mit DeepGreen: Gemeinsam den (grünen) Schatz heben. o-bib: das offene Bibliotheksjournal, 6 (1), 1-11. https://doi.org/10.5282/o-bib/2019H1S1-11

Many thanks for your attention

The presentation is available under a CC BY license: https://creativecommons.org/licenses/by/4.0/deed.de

All images are under License CC0: https://creativecommons.org/publicdomain/zero/1.0/deed.de

Please get back with questions to

Eike Wannick, Helmholtz Association, Helmholtz Open Science Coordination Office at the GFZ German Research Centre for Geosciences, Potsdam, Germany, eike.wannick@os.helmholtz.de

Julia Boltze, Zuse Institute Berlin – KOBV, Berlin, Germany, boltze@zib.de

Dr. Thomas Dierkes, Zuse Institute Berlin – KOBV, Berlin, Germany, dierkes@zib.de

Heinz Pampel, Helmholtz Association, Helmholtz Open Science Coordination Office at the GFZ German Research Centre for Geosciences, Potsdam, Germany, heinz.pampel@os.helmholtz.de

Markus Putnings, Friedrich-Alexander University Erlangen-Nürnberg (FAU), University Library, Erlangen, Germany, markus.putnings@fau.de

Website: http://deepgreen.kobv.de/

Prototype of the DeepGreen router: http://www.oa-deepgreen.de/