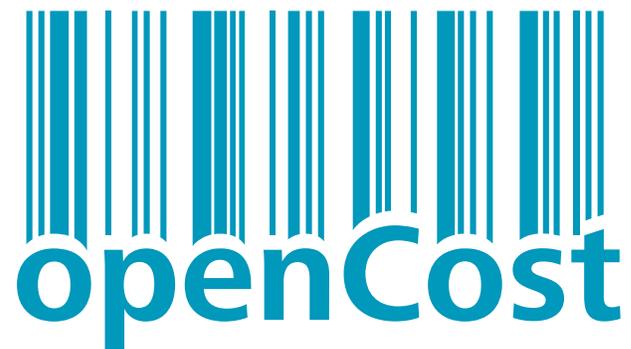


Bianca Schweighofer & Alexander Wagner (Eds.)



The Road to Publication
Cost Transparency

Proceedings of the Expert Workshop
Hamburg, October 05th – 07th 2022



Verlag Deutsches Elektronen-Synchrotron DESY
Hamburg
2023

Proceedings of the Expert Workshop
openCost —
The Road to Publication Cost Transparency



Hamburg, October 05th – 07th 2022

All presentations are available via

<https://indico.desy.de/event/35620>

Verlag Deutsches Elektronen-Synchrotron DESY
Notkestraße 85, 22607 Hamburg, Germany

[DESY-PROC-2023-01](#)

June 2023

ISSN 1435-8077

ISBN 978-3-945931-45-5

doi:[10.3204/PUBDB-2023-02961](https://doi.org/10.3204/PUBDB-2023-02961)



openCost is funded by grant [457354095](#).



This is an Open Access publication distributed under the terms of the [Creative Commons Attribution License 4.0](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Typeset by lua^BT_X

Publication cost transparency and the role of the Open Access Monitor Germany

Irene Barbers ¹

Forschungszentrum Jülich GmbH, Central Library

DOI: [10.3204/DESY-PROC-2023-01/10](https://doi.org/10.3204/DESY-PROC-2023-01/10)

Abstract

The Open Access Monitor Germany (OAM) records the publication output of German academic institutions in scientific journals and offers a freely available tool for the analysis of the aggregated datasets to libraries, funders, and researchers. Through analyses of subscription fees and publication fees, the OAM helps to monitor and support the transition of the publishing system towards open access. With the existing OpenAPC integration, the OAM has already implemented basic functions for working with cost data. Gold OA and Hybrid OA publication fees are displayed in the OAM interface, where additional grouping and representation options to those in OpenAPC are offered. As a next step towards a more complete cost transparency, the OAM will benefit from the data exchange enabled by openCost. In addition to the fees mentioned above, the OAM will collect via OpenAPC, and in return provide, additional cost data for example on color charges or page charges.

¹<https://orcid.org/0000-0003-2011-7444>

The second part of the presentation describes the role of the OAM in relation to funders. The OAM offers support to institutions applying for funding in the DFG's new Open Access Publication Funding program. By providing a specific filter set for the journal portfolios covered by transformative agreements and a curated list of open access journals that meet the DFG's funding criteria, the OAM enables institutions to collect the data required for their applications. At the same time, the OAM team is responsible for monitoring the publication output from participating institutions. We are building a dedicated database for the monitoring of the program's output and related costs, and establish a yearly reporting to the DFG. The Open Access Monitor Germany will ingest the data from the monitoring database if participating institutions are agreeable and will offer ready-to-use analyses for the whole program but also on the institutional level. The monitoring data will be in turn be delivered to OpenAPC.

10.1 Introduction

The Open Access Monitor Germany [137] is an openly available tool that aims to cover the full extent of publications published by German academic institutions in scholarly journals. By analysing subscription expenditure and publication costs, we observe and support the transition to Open Access in German scholarly publishing. The Open Access Monitor Germany (OAM) builds upon source systems that are already in existence or are currently in development [42].

With its connection to OpenAPC [53] the OAM has already implemented basic functions for working with cost data. Gold Open Access and Hybrid Open Access publication fees supplied by OpenAPC are on display in the OAM interface, with additional grouping and representation options to those in the OpenAPC web service. As a next step towards a more complete cost transparency, the OAM will benefit from the data exchange enabled by the openCost [37] project. In addition to the fees mentioned above, the OAM will collect via OpenAPC, and in return provide, additional cost data for example on colour charges or page charges.

The OAM offers support to institutions applying for funding in the new Open Access Publication Funding program [50] initiated by the German Research Foundation (DFG) [138]. By providing a specific filter set for the journal portfolios covered by transformative agreements and a curated list of Open Access journals that meet the DFG's funding criteria, the OAM enables institutions to collect the data required for their applications. At the same time, the OAM team is responsible for monitoring the publication output from participating institutions. We are building a dedicated database for the monitoring of the program's output

and related costs, and provide a regular reporting to the DFG. The OAM will ingest the data from the monitoring database if participating institutions are agreeable and will offer ready-to-use analyses for the whole program but also on the institutional level. The monitoring data will in turn be supplied to OpenAPC, again provided the institutions give their consent.

The following description of the OAM is in some parts based on two publications that describe the project, and our product. The *Serials Review* article [125] gives a general overview and outlines some use cases, whereas the article published in *Libreas* [139] (in German language) takes a more technical focus.

10.2 Source systems and production workflow in the OAM

To accomplish our aim of covering the full extent of German publications, we need a mechanism to assemble a pool of data. Our approach for harvesting such data is however not to collect in parallel, or repeatedly, reports from multiple providers such as single institutions or different publishers. Such a collection would be much too diverse, and both collecting and assembling would present an almost impossible effort in terms of managing different harvesting protocols and of standardization for example of affiliations or of other metadata. Instead, the principle behind the OAM is the current aggregation of data coming from few existing source systems.

Figure 10.1 shows the current and future source systems used to build the OAM database. Besides Unpaywall [140], several publication and citation databases, namely Dimensions [141], Web of Science [142], and Scopus [143], supply a weekly data feed. OpenAPC provides information on publication fees, whereas LAS:eR [144] and potentially, the ERM modules of Alma [145] and FOLIO [146] are sources for subscription fees. In the case of publication and citation data sources, we are conscious about the dependence on commercial providers as well as their possible shortcomings in terms of selection and bias. Therefore, we are always looking for new and preferably open sources. The integration of OpenAlex [147] is one of our next tasks in the project, envisaged for the year 2023.

To generate a comprehensive database (see Figure 10.2) from all the source systems mentioned, Unpaywall is our primary source. From there we collect article metadata such as DOI, journal, publisher, publication date and OA model for all publications with the document type “journal article”. Affiliation information and citation data come from Dimensions, Web of Science, and Scopus, who by

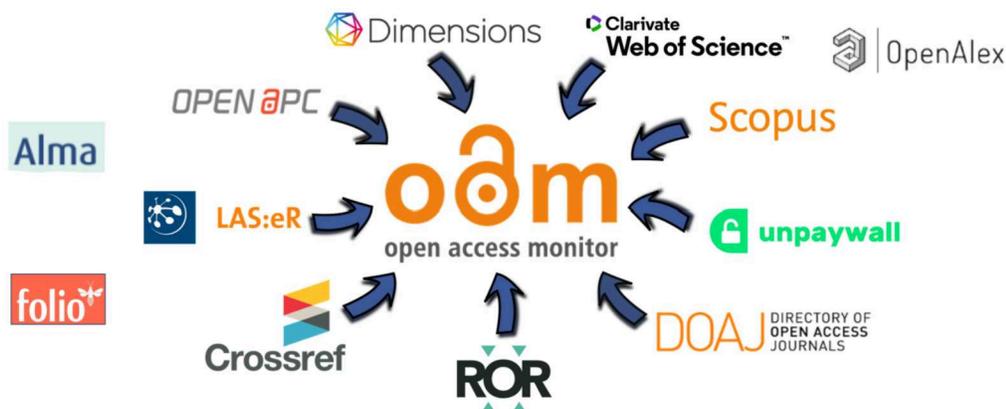


Figure 10.1: The Open Access Monitor Source Systems

agreement do not allow us to display their complete metadata sets. However, in addition to affiliations and citation counts, we display links to the respective data source on publication level.

A complex system for matching data via DOI and ROR-IDs helps to create connections to affiliations and citations for each article. Crossref [148] and DOAJ [9] are our sources for persistent identifiers and metadata on journal level, while the ROR registry [135] for organization identifiers serves as a supplier for controlled organization names across all sources. One of the major work packages within the project is the allocation of ROR-IDs to Web of Science and Scopus organization data, whereas Dimensions uses GRID-IDs [149], matching almost seamlessly to ROR.

Information on cost data via OpenAPC connects via a quite simple process of matching articles via DOI and institutions via ROR-ID. The integration of subscription fees via LAS:eR is a much more complex procedure, where we use direct API calls on the institutional LAS:eR accounts – by and with consent of the participating institutions - to harvest data on subscriptions, publishers, journal packages, and cost data for both local and consortial subscriptions. Due to the confidential nature of the information, institutional subscription analysis is the only part of the OAM tool that is exclusively accessible via login. Nevertheless, at the point where we will have a sufficient amount of data available, aggregations of subscription fees on national or publisher level will be made openly available, while keeping the institutions anonymous.

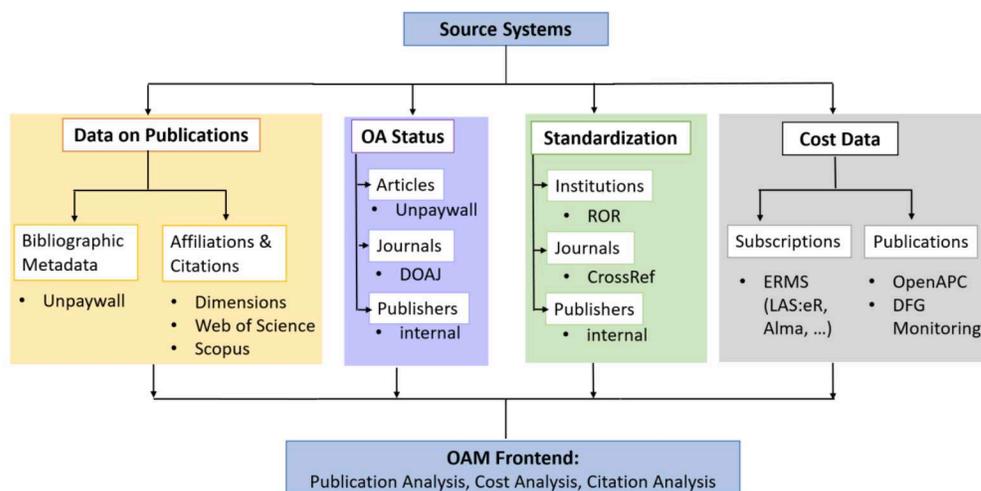


Figure 10.2: The OAM Database

10.3 Support for grant applications

The use of the OAM web application is described in detail in [125]. However, in the context of publication cost transparency, one specific use case within the publication analysis tool is of interest here. The collection of supportive data for grant applications plays an important role and is highly relevant for cost management and budget planning in libraries.

With the announcement of the DFG's new Open Access Publication Funding program [50] a new field of action opened up for the OAM's services. Because many institutions do not yet have adequate infrastructures or processes in place to record data on their own publication output, the OAM offers support to those institutions in performing the necessary data analyses. By providing a specific filter for the journal portfolios covered by transformative agreements and a curated list of Open Access journals that meet the DFG's funding criteria (quality-assured Open Access publication channels), the OAM enables institutions to gather the data required for their applications in an efficient way [125].

Part of the grant application is to set out in separate tables the eligible publications for the years 2019-2021 (proposals in 2022), or the years 2020-2022 (proposals in 2023), and distinguish between Gold Open Access and publications based on transformative agreements or other Open Access models. To support this use case, we provide filters including those journals that are part of transformative agreements. In accordance with the DFG funding program [50], we use only the agreements registered with the ESAC [136] initiative.

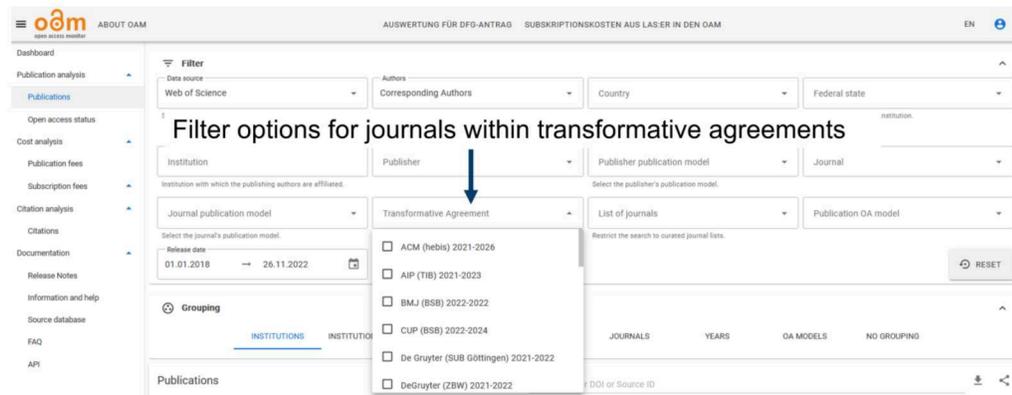


Figure 10.3: Filter options for journals within transformative agreements in the OAM user interface

OAM users can choose the agreements their institution participate with as shown in [Figure 10.3](#) and as a result receive the publication output they have had or could have had by Corresponding Authors within these agreements in the past. With this data as a basis, they can calculate the potential publication output in the future and then use the outcome for the grant application in the DFG program Open Access Publication Costs.

The same principle works for Gold Open Access journals that are eligible within the DFG funding program. To retrieve the necessary data for the DFG grant application, users can choose their institution and the list “DFG-Anträge”.

These supporting features for grant applications lead to the central topic of publication cost transparency. On the base of the monitoring process installed for the DFG funding program, the cost information on funded publications will be visible within the OAM and connected systems.

10.4 Publication cost transparency

To enable the monitoring of the DFG Funding Program, operated by the Central Library of Forschungszentrum Jülich on behalf of the DFG, a standardized reporting data schema was developed and has been made available to all participating institutions. In the first cycle, there are 75 institutions participating in the program who will receive funding for the period of 2022 to 2024. Within the program, the yearly reporting of detailed extensive publication and cost data is a mandatory requirement.

Table 10.1 provides a simplified overview, demonstrating the essentials of the data schema. In order to achieve a good level of standardization, there are sets of fixed properties for the data elements to choose from wherever possible, for example in the fields provided for the license, the publication type, or the allocation of a transformative agreement. Additional tables within the metadata schema facilitate the reporting of lump payments for transformative agreements or memberships. As a special feature, the metadata schema provides the possibility to report costs that are not eligible for funding. Reporting those costs is not mandatory, but strongly encouraged to help establish a complete and transparent monitoring, and an infrastructure for the participating institutions' own monitoring efforts.

Data Element		Description
DOI	mandatory	In case of several cost types relating to the same publication, use a separate row for each cost type, and repeat DOI
Name of Publisher, Server, Repository	mandatory	
Publication Type	mandatory	Choose from dropdown journal article; book; book part; conference proceedings; dataset; preprint; software; other
CC-License	mandatory	Choose from dropdown CC BY; CC BY-SA; CC BY-ND; CC BY-NC; CC BY-NC-SA; CC BY-NC-ND; Other; none
Currency	mandatory	Example: USD, GBP, or other
Amount in original currency	mandatory	Set to 0.00 for articles from transformative agreements or from membership agreements. Use transformative agreements or membership tabs to record central payment.
Net amount (Euro)	mandatory	Set to 0.00 for articles from transformative agreements or from membership agreements. Use transformative agreements or membership tabs to record central payment.
Tax value	mandatory	Applicable tax value or reverse charge value
Total amount (euro)	mandatory	Set to 0.00 for articles from transformative agreements or from membership agreements. Use transformative agreements or membership tabs to record central payment.

Data Element		Description
Funding amount DFG	mandatory	Set to 0.00 for articles from transformative agreements or from membership agreements. Use transformative agreements or membership tabs to record central to central payment.
Cost type (not eligible for funding)	optional	Choose from dropdown Colour charge; Cover charge; Hybrid OA-Fee; Page Charge; Publication Charge; Reprint; Submission fee; Other
Membership	optional	Name of membership agreement; use tab for memberships to record payments
Transformative agreement	mandatory; can be set to “none”	Name of transformative agreement; use tab for transformative agreements to record payments Choose from dropdown none; ACM (hebis); AIP (TIB); BMJ (BSB); CUP (BSB); DeGruyter (SUB Göttingen); DeGruyter (ZBW); ECS (TIB); Hogrefe (SUB Göttingen); IOP (TIB); Karger (BSB); Nature (MPDL); RSC (TIB); Sage (BSB); SPIE (TIB); Springer (DEAL); Thieme 1 (ZB MED/FZJ); Thieme 2 (ZB MED/FZJ); Wiley (DEAL)
Funding year	mandatory	Year in which funding was received from DFG Leave blank for articles from transformative agreements or from membership agreements
Invoicing year	mandatory	Year of publisher's invoice Leave blank for articles from transformative agreements or from membership agreements
DFG Grant number	optional	Grant number of DFG research funding the publication has originated from Not mandatory at this stage, but relevant for the second proposal phase
DFG subject area	mandatory	Choose from dropdown Social Sciences and Humanities; Life Sciences; Science; Engineering; Multidisciplinary

Table 10.1: Overview of the monitoring metadata schema for the DFG OA Funding Program

There will be a dedicated database built for the monitoring of the project, which will enable evaluation reports to be provided to the DFG on a yearly basis.

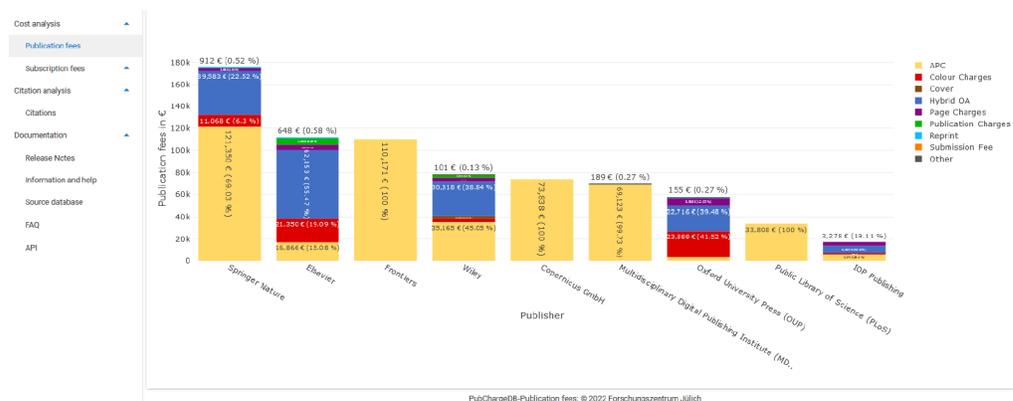


Figure 10.4: Institutional publication costs broken down by cost type and publisher in the OAM user interface

Additionally, libraries not participating in the funding program can also benefit by using the metadata schema for their own cost monitoring projects. Finally, the OAM will be able to use the same data if participating institutions are willing to share and will offer ready-to-use analyses and graphs on institutional but also on national level.

10.5 Publication cost visualizations

Publication cost data collected by the OAM via DFG Monitoring and/or OpenAPC will be available for evaluation in the cost analysis area of the web application. To illustrate this, [Figure 10.5](#) shows a preliminary view of a visualization for publication costs broken down by type and publisher. The OAM data model has already been adapted to be able to receive such granular publication cost data and will synchronize with the data model developed for OpenAPC.

[Figure 10.6](#) is a view of subscription cost integration displayed in contrast to an institution's publications spend. For the time being, the subscription cost information within the OAM depicts the LAS:eR structure, and is therefore not very granular yet. There is for example no differentiation between the publish and the read shares of subscription costs because this is not yet specified as a flag on the LAS:eR side.

Information on licenses and subscription fees can complete the picture both on the level of individual institutions and, ultimately, on a national level. Ideally,

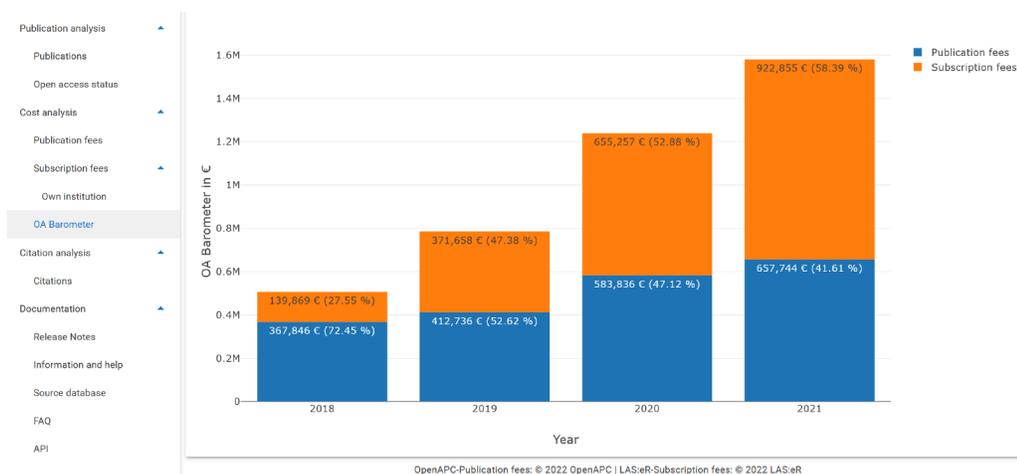


Figure 10.5: Institutional subscription fees vs. publication fees broken down by year.

institutions will in the future receive a complete overview of their portfolio of licenses, publications, and the associated fees to support them in managing information on their income, and on their expenditures for scientific publications. The goal is to provide each participating institution with data and visualizations like the Jülich Open Access Barometer [88], which in turn contributes to the concept of the information budget.

10.6 Conclusions and Outlook

The DFG Monitoring Project will receive data from participating institutions on a mandatory basis. The same data will then be made available to OpenAPC and the OAM, if the institutions are agreeable. This saves time and trouble for participating institutions who will not need to prepare and supply the same data several times over. Institutions that do not participate in the DFG program can supply their cost data directly to OpenAPC thanks to the published metadata format and in this way can benefit from the services of both OpenAPC and the OAM

The OAM constitutes a central part in the landscape of the total cost of publication transparency initiatives, working together with OpenAPC and the DFG program monitoring; and enabling and further developing such transparency

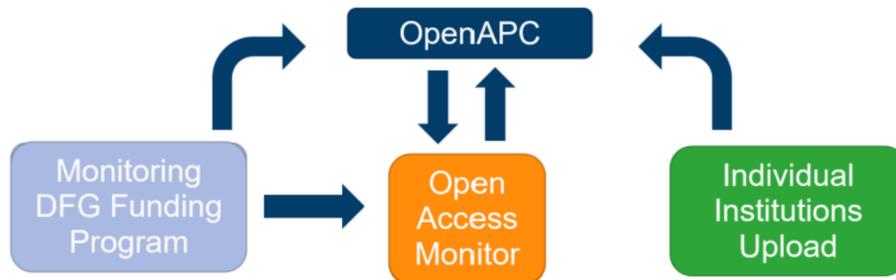


Figure 10.6: Data sharing cycle for total publication cost transparency

through the collaboration with projects like openCost with a technical focus, and Transform2Open [49] with a focus on workflows and competencies. These contributions will support individual institutions in their aim to establish an information budget [84] whereas on a broader level, the OAM continues to pursue the goal of observing and supporting the transition of the publication system towards Open Access, with a reinforced focus on cost transparency.

10.7 Acknowledgements

The German Federal Ministry of Education and Research (BMBF) supports the Open Access Monitor Germany, and this publication, under Grant FKZ 16OAMO001.

References

- [9] *Directory of Open Access Journals. DOAJ*. en. URL: <https://doaj.org/> (visited on 2022-12-06) (cit. on pp. 3, 107, 116).
- [37] openCost. *openCost: Projekt*. URL: <https://www.opencost.de/projekt/> (visited on 2022-11-27) (cit. on pp. 14, 16, 114).
- [42] Zentralbibliothek Forschungszentrum Jülich, ed. *Open Access Monitor project website*. URL: https://go.fzj.de/OAM_en (visited on 2022-11-27) (cit. on pp. 16, 114).

- [49] Roland Bertelmann et al. *Transform2Open. Kostenmonitoring, Kriterien, Kompetenzen und Prozesse der Open-Access-Transformation*. de. Helmholtz Open Science Office, 2022. 22 pp. DOI: [10.48440/OS.HELMHOLTZ.054](https://doi.org/10.48440/OS.HELMHOLTZ.054) (cit. on pp. 21, 46, 97, 100, 123).
- [50] Deutsche Forschungsgemeinschaft. *Open Access Publication Funding: Guidelines and Supplementary Instructions*. DFG form 12.21. 2022 (cit. on pp. 23, 114, 117).
- [53] Bielefeld University Library and DINI working group Electronic Publishing. *OpenAPC Home*. URL: <https://openapc.net/> (visited on 2022-11-27) (cit. on pp. 32, 114).
- [84] Heinz Pampel. “From library budget to information budget: fostering transparency in the transformation towards open access”. In: *Insights the UKSG journal* 35 (2022). DOI: [10.1629/uksg.576](https://doi.org/10.1629/uksg.576) (cit. on pp. 47, 77, 96, 97, 123).
- [88] Zentralbibliothek Forschungszentrum Jülich. *Open Access Barometer*. URL: <https://www.fz-juelich.de/en/zb/open-science/open-access/oa-barometer> (visited on 2022-12-13) (cit. on pp. 54, 122).
- [125] Irene Barbers, Franziska Stanzel, and Bernhard Mittermaier. “Open Access Monitor Germany: Best Practice in Providing Metrics for Analysis and Decision-Making”. In: *Serials Review* 48.1-2 (2022), pp. 49–62. DOI: [10.1080/00987913.2022.2066968](https://doi.org/10.1080/00987913.2022.2066968) (cit. on pp. 99, 115, 117).
- [135] *Research Organization Registry (ROR)*. URL: <https://ror.org/> (visited on 2022-12-08) (cit. on pp. 108, 116).
- [136] ESAC Community. *ESAC Transformative Agreement Registry*. en. URL: <https://esac-initiative.org/about/transformative-agreements/agreement-registry/> (visited on 2022-12-06) (cit. on pp. 110, 117).
- [137] Zentralbibliothek Forschungszentrum Jülich, ed. *Open Access Monitor web application*. URL: <https://open-access-monitor.de/?culture=en> (visited on 2022-11-27) (cit. on p. 114).
- [138] Deutsche Forschungsgemeinschaft. *DFG. Deutsche Forschungsgemeinschaft*. URL: <https://www.dfg.de/en/index.jsp> (visited on 2022-11-27) (cit. on p. 114).

- [139] Franziska Stanzel et al. “Big Scholarly Data im Open Access Monitor: ein Werkstattbericht”. In: *Libreas. Library Ideas* 41 (2022), pp. 1–20. DOI: [10.18452/24797](https://doi.org/10.18452/24797) (cit. on p. 115).
- [140] OurResearch. *Unpaywall: Data feed*. URL: <https://unpaywall.org/products/data-feed> (visited on 2022-11-27) (cit. on p. 115).
- [141] Digital Science & Research Solutions Ltd. *Dimensions*. URL: <https://www.dimensions.ai/> (visited on 2022-11-27) (cit. on p. 115).
- [142] Clarivate. *Web of Science*. URL: <https://www.webofscience.com/wos/woscc/basic-search> (visited on 2022-11-27) (cit. on p. 115).
- [143] Elsevier B.V. *Scopus*. URL: <https://www.scopus.com> (visited on 2022-11-27) (cit. on p. 115).
- [144] Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz). *LAS:eR*. URL: <https://laser.hbz-nrw.de/> (visited on 2022-11-27) (cit. on p. 115).
- [145] ExLibris. *Alma*. URL: <https://www.exlibrisgroup.com/products/alma-library-services-platform/> (visited on 2022-11-27) (cit. on p. 115).
- [146] The Open Library Foundation. *FOLIO*. URL: <https://www.folio.org/> (visited on 2022-11-27) (cit. on p. 115).
- [147] OurResearch. *OpenAlex*. URL: <https://openalex.org/> (visited on 2022-11-27) (cit. on p. 115).
- [148] *Crossref*. URL: <https://www.crossref.org/> (visited on 2022-11-27) (cit. on p. 116).
- [149] Digital Science & Research Solutions Ltd. *GRID*. URL: <https://grid.ac/> (visited on 2022-11-27) (cit. on p. 116).