Bianca Schweighofer & Alexander Wagner (Eds.)

The Road to Publication Cost Transparency

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

Verlag Deutsches Elektronen-Synchrotron DESY
Hamburg
2023

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

Hamburg, October 05^{th} – 07^{th} 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*

Deutsche Forschungsgemeinschaft

openCost is funded by grant 457354095.

Typeset by luaLaTeX

Contents

About the Project

From October 5th to 7th, the international expert workshop openCost: the road to publication cost transparency took place at the Deutsches Elektronen-Synchrotron, DESY in Hamburg. There the project partners DESY, UB Bielefeld and UB Regensburg presented first project results. The workshop also served as a platform for knowledge exchange between national and international experts. Speakers from eight different countries contributed their perspectives.

First day: Introduction and knowledge exchange with experts

On the first day of the workshop, the project partners introduced themselves in the block The openCost Project and addressed their respective roles in the project. Afterwards, Bernhard Mittermaier, Head of the Central Library at Forschungszentrum Jülich, gave a keynote on the topic of Information Budget. He provided important impulses for the workshop. In the thematic block Knowledge exchange between national and international experts, experts from the California Digital Library and Sikt (Norway) reported on their experiences with publication costs and cost recording.

Second day: Knowledge exchange and interactive hands-on labs

The second day of the workshop continued from this point. International and national experts from Jisc, AT2OA2, the National Library of Finland and the DFG project Transform2Open contributed their respective perspectives on publication costs. Afterwards, the 40 participants had the opportunity to work together on a metadata schema in various hands-on labs.

In the hands-on lab Metadata schema, participants discussed the current internal proposal for a metadata schema. The participants provided fruitful suggestions.

The hands-on lab Terminology was dedicated to clear, unambiguous definitions of terms to describe cost data. And the third hands-on lab dealt with the consideration of payments that cannot be clearly assigned to a publication. In particular, the participants addressed publishing agreements, cost sharing, and articles without DOI. Finally, the day ended after two more presentations from the last thematic block, New Services and Subsequent Use. Experts from the Electronic Journals Library (EZB) and Unsub/OpenAlex reported on how openCost data could be used in other services in the future.

Third day: Wrap-up and subsequent use

On the last day of the workshop, the lab results from the previous day were summarized for all participants and lively discussed. These results will serve as a starting point to enhance and fine tune the project team's current internal metadata schema. Afterwards, experts from the Open Access Monitor Germany and OA Switchboard rounded off the workshop with two more interesting presentations from the block New Services and Subsequent Use.

The expert workshop summarizes the desiderata of the individual participants and the committees they represent. The present conference proceedings document the majority of the presentations given at the workshop. The results of the workshop, especially those of the hands-on labs, have been incorporated into the metadata schema that will also be presented in these proceedings.

1

Using the institutional repository to store data related to payments

Gernot Deinzer[1]
Universität Regensburg, Universitätsbibliothek

DOI: 10.3204/DESY-PROC-2023-01/01

Abstract

At the University of Regensburg, all payments for publications must be made by the library. The fees, invoices and additional payment information are stored together with the document in the institutional repository. We present the motivation for this decision, the related workflow, and benefits from storing data in the repository.

1.1 Introduction

Open access has become a successful business model publishing scholarly information and is adapted by nearly all publishing houses around the world. By increasing the number of open access publications one immediately asks about the prices keeping in mind the price increase of serials. At least one should know about the spendings for open access and publishing in general. Even this is not a very easy task, because publishing fees are in the normal case not addressed to the library or the university, but to the corresponding author himself.

[1] https://orcid.org/0000-0002-7462-3847

1.2 Preliminary considerations

Looking at the requirements from funders (see e. g. [1]) politics (see e. g. [2]), executive boards of universities and other related stakeholders, universities have to deliver a huge amount of numbers related to open access. These cannot easily be evaluated due to the fact, that not all numbers are stored in the same place and searches over platforms must be performed with the risk of losing some data.

Typical questions regarding open access are

- How many publications are open access?
- Why are publications open access?
- How is the quality of open access publications assured?
- What is the total spending for open access publications?
- What is the amount paid by central funding from the library?
- Can data be reused by publishing with suitable licenses?

These questions can be asked on an institutional level as well as over different institutions e. g. federal states. To answer these questions, one must either collect the data on the level of the whole institutions and to provide these data in a standardized way to allow third parties to collect and gather the respective information.

1.3 Open access situation

There are a lot of possible reasons why a publication is open access and how we are spending resources for it. This can be either be money or staff cost. Mainly we distinguish four possible ways:

- **Primary open access publications at publisher site**

  This involves a payment for each publication published open access. These can be payment for article in open access journals[2] or buying a free license

[2] Gold Open Access journals are journals where the complete content is published open Access.

to an article in a closed journal[3]. Books can often be published openly by paying a so-called book processing charge. But even in contracts with publisher payments for each article can be included. These are the so-called transformative agreements. Best examples are the German DEAL [3] contracts, where each article is assigned with Publish and Read Fees. Transformative agreements with a Read and Publish model[4] we would classify herein, because even if the contract is based on a subscription fee, all or a certain number of open access publications are included in the total fee.

- **Support of community-based business models**

  We pay a specific amount for a certain period and all publications within this are open access. Open Library of Humities (OLH) [4], SciPost [5], SCOAP³ [6], PLOS Community Action Publishing [7], are well known examples for such a business model. But also, pledging models like KOALA [8] supporting of infrastructures like the DOAJ [9] we would allocate in this category.

- **Parallel Publication**

  Further on we consult our researchers making their publications open access available in the institutional (or a subject based) repository. Even if we don't spend money directly for the publication, we invest in library staff. They consult the researchers by answering legal questions and support assuring publishing rights in agreements with publishers.

- **Institutional offers**

  Beside this we provide an infrastructure to publish at the university. Within this infrastructure we offer a service as a publisher, where the primary publication is both open access and printed. For publishing journals open access we run a platform at the university. Theses, research data, and software can be published in the institutional repository. We don't charge the authors or editors any fee, even if we investigate in staff and the technical infrastructure, e. g. servers and storage.

[3] Journals with the choice of closed and open publishing are known as hybrid journals.

[4] Read and Publishing business model means, that the payment is based on the subscription fee and a possible additional fee for publications with an open license.

1.4 Information budget

The information budget [10, 11] is defined as the whole spendings for information material such as licenses to journals and eBooks, buying printed copies of journals or books, paying fees for publishing, providing inter library loan or pay per view. Just at the first glance one sees immediately that it is very hard to identify all payments regarding this budget. A lot of payments are directly done by authors either via the university administration or even worse by themselves. We focus ourselves here on the payments directly for publishing. All other payments (mainly for reading) are much easier to determine. We assume that all the acquisition and licensing of information material is done by the library and so all the numbers are well known. Therefore we have a closer look to all the different parts related to publishing. First, we have the classical payments for a single article. This can be an APC for publishing the article open access either in a Gold Open Access journal or a hybrid journal, charges applied for different reasons like color images (often referred as color charges), due to the length of the article (page charges) or even additional services like processing it to the peer review (submitting charges), printing it on the cover (cover charges) or ordering author copies. There is no difference for articles in conference proceedings. Even for books there exists nowadays, besides the classical author charges for printing a copy, new charges for publishing open access. These payments are hardly to identify in the accounting system of a university.

In an ideal world all invoices for information would be covered by the library and thus easily identifiable. But the real world differs. Most libraries are not paying for all information especially for the charges related to publishing. There are e. g. charges which are not covered by the library because they do not fulfill the funding requirements. Often the hybrid APCs are excluded due to the so-called double dipping issue. Although payments not related to open access are not fundable.

Often open access fees like APCs are just known to certain range, because requirements are here too strict to apply like a certain maximal threshold. Sometimes authors just don't apply for a funding. This might be the case where funding grants still must be spent. In real life we have therefore a situation that charges related to publications are paid from different sites: publication fees are paid either by the library, by the researcher from their own budget or by funding bodies like the Deutsche Forschungsgemeinschaft (DFG).

1.5 Central Invoicing

To get an overall insight, how much was paid for publications at the university, we had a look at the process of buying printed books and compared it with the workflow for paying publication fees. In both cases there is the possibility of saving money due to certain agreements or contracts with publisher or book sellers. The workload will be decreased due to gathering invoices on a regular base. And often these are the same contract partners, the publisher. The question was: why is there a centralized process for buying books and why does no equivalent workflow exist for publication fees?

The simple answer to this question is the inventory of the printed book. Each book gets a stamp that it is owned by the university. In an electronic publication this is not possible and thus the publication has not to pass the central office with the stamp, in particular the library.

At the beginning of the year 2021 a change was introduced. The process of paying publication fees followed the workflow of buying books. The executive board of the university decided that all invoices regarding payment for publication must pass the library. This was done due to the recommendations[5] of both the German Rectors' conference (HRK) and the library committee of the university. At the same time the administration of the university introduced the central billing invoice for all invoices of the university. This includes the publication fees. A new process for paying publication fees was installed. The invoices must be sent to the central billing office directly by the publishers and are then assigned to the library. This is done either by stating the library in the billing address or automatically if there are certain key words in the invoice (like APC, page charges or color charges) or if the billing partner is a known publisher. Afterwards the library gets in contact with the author and clarifies how to pay the invoice. This can be done from the central publication fund, from projects founded by funders like the DFG or by the resources of the researchers. Even a splitting of the invoice over different resources is possible. The library pays the invoice and makes all the internal money transfer. A copy of the publication will be stored in the institutional repository. Open access if the publication was published with a reusable right or with restricted access if the publication is closed access or has no reusable rights assigned. This corresponds to the inventory of a book.

[5] Rundschreiben Nr. 018/2020 ("Fortführung es DEAL-Vertrags mit Springer Nature") vom 30.06.2020

1.6 Metadata of publication

If we are looking at the metadata of a publication, we can divide it in several distinct parts [12]

- Bibliographic metadata: all the data describing the publication like author, title, keywords, abstract
- Technical metadata: technical information about the publication like file-type, filesize, checksums but also the dates (submitting, accepting, publication)
- Legal metadata: this are mainly the rights assigned to the publication like the licenses
- Organizational metadata: this includes all the relationship of a publication with an organization like a research institute, the groups therein, projects, and funders.

These are the well-known metadata, and they are recommended to be stored for each publication in repositories. But if we are broadening our view on publications and with all the new publishing models, we can enlarge the set of metadata with information about payments. These can include payments directly for the publishing like APCs, page charges etc. or information about the publication being part of a contract which allows the article to be published open access or to give the article distinct reuse options. Best known examples for such contract are the so-called transformative agreements like DEAL, memberships like SCOAP³ or consortia models like OLH or SciPost

1.7 Adding metadata to a publication

For every publication it must be clear, whether it is open access or not and, if it is open access, why it is open access [13]. Therefore, first of all we specify why an article is open access or if there are no open access possibilities for it. This can be either the information that it was paid on a single article level (APC in gold or hybrid journal with distinction paid by the University of Regensburg or another research institute), no payment due to the business model of the journal (Gold OA journals without APCs, Diamond OA), part of a membership or transformative agreement or parallel publication due to some transferred rights (German copyright

law, policies of publisher). Further on, if a payment was done by the university, the details of the payment are also added to the metadata.

These are the total costs, costs related to open access, the part paid by central funding as well as the date of payment and the internal transaction numbers for each payment. The information of the invoice is added to the metadata. This consists of the invoice date, the amount and currency printed on the invoice, the invoice number and if possible, the article reference number. We also specify if there is an acknowledgement to a funder or the reference to research data as these are sometimes prerequisites for a funding. With this information we can connect the open access and payment information of a publication to the publication itself from authors from the university.

1.8 Gathering information

To get the information even outside the institutional repository, one must connect the different blocks of metadata. To establish a connection, one needs a unique and persistent identifier. For published articles often this identifier exists in form of a DOI and everything can be linked correctly. But for non-published articles or books, no such identifier exists, and the linking is extremely difficult. But as long as DOIs are crucial for a long-term citation of a digital object, each scholarly publication independently if it is an journal article, a book, a book section or a conference proceeding should have an assigned DOI. If a publication has more than one DOI, e. g. because the preprint, the publications and a parallel version in a repository have a distinct DOI, on should establish a connection of the different DOIs like is preprint of, is equivalent to, is version of etc. It makes in our opinion completely sense to assign DOIs to digital copies in a repository as far as the DOI points to exact one digital object and not to equivalent class of objects.

1.9 Displaying the information of open access

A big advantage to have all the data (organizational, bibliographic, technical, legal information, payments) in one place (the institutional repository) is the possibility to analyze the data in depth. No limits are set to combine all the information and a broad overview over publishing activities can be generated. This includes e. g. costs and business models for different subjects, distribution of costs within the

research institution. Money spent for different publishers, relationship between different open access business models and subjects and so on.

Even reports for executive boards and funders can easily and automatically generated. This is extremely useful to argue for central funding of fees related to open access as well as to comply with funding requirements, e. g. for the DFG-program Open-Access-Publikationskosten [14].

1.10 Exchanging data

Of special interest is the aggregation of data over several institutions. This has already been done for APCs in the project OpenAPC [15]. To our knowledge this is the first initiative to collect data about publication charges and gather the information in one place. To participate one must deliver the data in a special form defined by the project. The restriction to APCs leads to a very small set of data to be provided. By extending this set to all payments for publishing, which is not an easy task, new possibilities to analyze payments and money flows can be generated. But this data must be provided with an open license like CC0 and in a standardized and applicable way. Metadata is usually exchanged between different systems with special xml-schemes, which are provided over a common interface the so called OAI-Interface [16]. Even for payment information the way can and should be used as service providers can easily harvest the data from the different institutions (data providers) and build upon this data new services beginning with large databases ending with different statistical tools. Even the mechanism of constraining to specific sets, would allow stakeholders to get just the information they are interested in. Up to now no schema for payments analogous to well established schemes for bibliographic metadata like Dublin Core [17] or Datacite [18] has been defined.

1.11 Summary

At the University of Regensburg we established a centralized workflow to process payments for publishing. All the payment information is stored as metadata related to a publication in the institutional repository. Even this is combined in the first step with additional work in the open access team, the benefits in generating automated reports and statistics is enormous. By storing the data in a structured way in in open accessible platform the data is available for anyone who

is interested in processing cost data. New evaluations of payments to publishers are possible if libraries worldwide provide this information in a standardized way. Stakeholders like funding bodies can get a deep insight about publication spending over different institutions.

References

[1] Forschungszentrum Jülich. *Monitoring DFG-Förderprogramm „Open Access Publikationskosten". Erläuterungen zum Datenschema*. Apr. 6, 2023. URL: https://www.fz-juelich.de/de/zb/downloads/open-science/monitoring-oa-publikationskosten-datenschema-erlaeuterungen-2023-01-25.pdf/@@download/file (cit. on pp. 2, 16, 23).

[2] Maxi Kindling et al. "Open-Access-Landesstrategien und der Weg zur offenen Wissenschaft – Möglichkeiten und Herausforderungen am Beispiel Berlin". de. In: (2021). DOI: 10.5281/ZENODO.5535754 (cit. on p. 2).

[3] DEAL Konsortium. *About DEAL*. Apr. 6, 2023. URL: https://deal-konsortium.de/en/about-deal (cit. on p. 3).

[4] Open Library of Humanities. *Open Library of Humanities*. Apr. 6, 2023. URL: https://www.openlibhums.org/ (cit. on p. 3).

[5] SciPost Foundation. *SciPost*. Apr. 6, 2023. URL: https://scipost.org/ (cit. on p. 3).

[6] CERN. *Sponsoring Consortium for Open Access Publishing in Particle Physics. SCOAP³*. Apr. 6, 2023. URL: https://scoap3.org (cit. on pp. 3, 16).

[7] Public Library of Science. *PLOS Community Action Publishing*. Apr. 6, 2023. URL: https://plos.org/resources/community-action-publishing/ (cit. on p. 3).

[8] TIB Leibnitz Centre for Science and Technology University Library. *Konsortiale Open-Access-Lösungen aufbauen. KOALA*. Apr. 6, 2023. URL: https://projects.tib.eu/koala/en/ (cit. on p. 3).

[9] *Directory of Open Access Journals. DOAJ*. en. URL: https://doaj.org/ (visited on 2022-12-06) (cit. on pp. 3, 105, 114).

[10] Bernhard Mittermaier. “Informationsbudget”. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 9.4 (2022), pp. 1–17. DOI: 10.5282/o-bib/5864 (cit. on pp. 4, 46).

[11] German Science and Humanities Council. “Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access”. de. In: (2022). DOI: 10.57674/FYRC-VB61 (cit. on p. 4).

[12] Heinz Pampel. *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten*. de. Arbeitspapier. 2019. DOI: 10.2312/OS.HELMHOLTZ.006 (cit. on pp. 6, 17, 47, 72, 75).

[13] DINI AG Elektronisches Publizieren. *Dokumentation der Rechtsgrundlage für Veröffentlichungen in Repositorienmetadaten. Positionspapier*. de. DINI Schriften 20. Humboldt-Universität zu Berlin, 2022. DOI: 10.18452/23191 (cit. on p. 6).

[14] Deutsche Forschungsgemeinschaft. *Funding Programme: “Open Access Publication Funding”*. URL: https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/open_access_publikationskosten/ (visited on 2022-12-13) (cit. on pp. 8, 16, 21, 22, 47, 49).

[15] Bielefeld University Library and DINI working group Electronic Publishing. *OpenAPC*. en. URL: https://openapc.net/ (visited on 2022-12-06) (cit. on pp. 8, 16, 103).

[16] Open Archives Initiative. *Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14*. Jan. 8, 2015. URL: http://www.openarchives.org/OAI/openarchivesprotocol.html (visited on 2023-04-06) (cit. on pp. 8, 14, 15, 19).

[17] Dublin Core™ Metadata Initiative. *XML schemas to support the Guidelines for implementing Dublin Core™ in XML*. Apr. 6, 2023. URL: https://www.dublincore.org/schemas/xmls/ (cit. on p. 8).

[18] DataCite. *DataCite Metadata Schema*. Apr. 6, 2023. URL: https://schema.datacite.org/ (cit. on p. 8).

2

JOIN² and openCost

Lisa-Marie Stein[1]
Deutsches Elektronen-Synchrotron DESY, Central Library

Alexander Wagner[2]
Deutsches Elektronen-Synchrotron DESY, Central Library

DOI: 10.3204/DESY-PROC-2023-01/02

Abstract

Founded more than ten years ago, the JOIN² collaboration [19] brings together eight research institutions for the development and operation of a full-fledged shared scholarly publication database and repository based on the Invenio open source framework for large-scale digital libraries. Preferring simplicity to complexity one of our corner stones is to build on well-defined workflows. This enabled JOIN² to extend its services from representing the scholarly output of the member institutions to serve as the Open Access repository on site and gradually add additional services. One of these is the management of publication costs data alongside the publications within the repository. From the start we headed to model all costs we came across in scholarly publishing, be it open or closed access, to further cost transparency. This naturally led JOIN² to be one of the core members of openCost, with the Deutsches Elektronen-Synchrotron DESY representing the collaboration.

[1] https://orcid.org/0000-0001-7905-0462
[2] https://orcid.org/0000-0001-9846-5516

2.1 Who is JOIN²?

JOIN² is an acronym for **J**UST AN**O**THER **I**NVENIO I**N**STANCE **TO JOIN**. As the spoken word suggests, the project invites to become part of a collaboration that jointly works on the challenges arising in practical life[3].

Originally, JOIN² [19] was founded in 2010 by the libraries of Deutsches Elektronen-Synchrotron DESY, Forschungszentrum Jülich and GSI Helmholtzzentrum für Schwerionenforschung to replace existing solutions for publications databases and repositories [20, 21]. Over the years the JOIN² collaboration grew, and beyond the initial partners now also comprises[4] Deutsches Krebsforschungszentrum DKFZ (Heidelberg), Deutsches Zentrum für Neurodegenerative Erkrankungen DZNE (Bonn), Maier-Leibnitz-Zentrum (Garching), Museum Zitadelle (Jülich)[5] and RWTH Aachen University [22]. Figure 2.1 gives a map of all partners and associated laboratory sites that use JOIN² in Germany.

The institutional repositories maintained by JOIN² [22, 23] aim to serve multiple requirements, combine different applications in one system, and offer easy-to-use solutions suitable for everyday business (see e. g. [24–27] for an overview). They are intended to handle any kind of scientific literature and strategically foster Open Science. In addition to the collection and exposition of the scientific results of an institution, functioning as a publications database [25], a protected workspace for scientists has also been created [28, 29]. Publications and other resources can be exchanged and shared for various purposes, and the attachment of full texts is explicitly encouraged. Structured exports allow the reuse of the bibliographic data for upcoming publications or feed the publication list on the web. The integration with literature management tools and complex output filters, which are essential for reuse in the publishing process, also serve evaluations and budgeting of the administration, all operated via one common system. A detailed and constant requirement review of the metadata as well as the established authority control [30, 31] are important preconditions. The high quality of metadata even enables some partners to run their library system within their JOIN² instance [32].

[3] Using a very pragmatic approach to solve the problems at hand, the collaboration intitially did not even have a name. This was only estbablished in 2015 as more parters joined the project.

[4] As a result of the war of Russia against Ukraine, the membership of the Joint Institute for Nuclear Research, JINR (Dubna, Russia) is suspended.

[5] This instance uses JOIN² only for their library catalogue and does not contribute to the subsequent discussion.

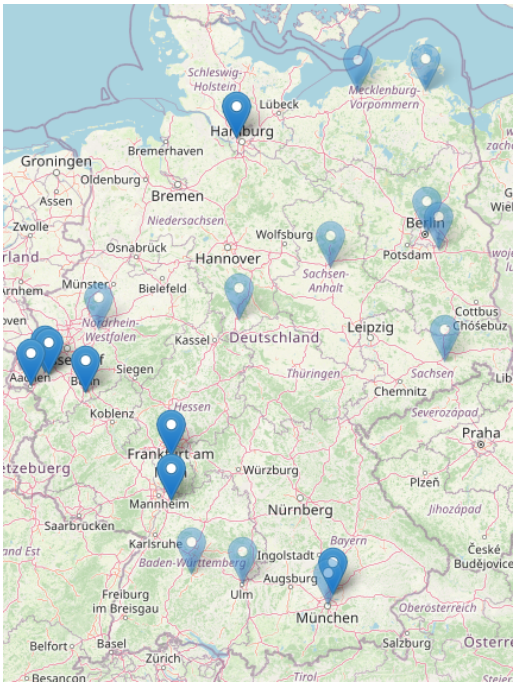

Figure 2.1: German partners of the JOIN² collaboration as of 2022: The dark dots mark locations where an instance is in operation, the shaded ones mark laboratory sites associated with the partners that also use one of these services. E. g. DESY operates PUBDB in Hamburg, but has a second laboratory in Zeuthen near Berlin.

In 2022 JOIN² provided service to more than 23.000 staff members. More than 6.000 visiting scientists also use the services e. g. to report their achievements. In total, the running instances provided more than 609.000 bibliographic records, where more than 98.000 also offer a full text available in Open Access. Additionally, if licencing restricts the supply of the full text, for a large number of articles the bibliographic records hold pointers to other repositories (e. g. Pubmed Central [33] or arXiv [34] etc.) where the full text is available.

Besides bibliographic data the JOIN² infrastructure shares nearly 147.000 authority records [30, 31] to ensure the high quality of the metadata provided plus a number of author authority records individual to each instance[6]. With

[6]To give an idea at DESY these are currently more than 44.000 additional records, while at RWTH Aachen University more than 210.000 internal authorities are required.

the sole exception of the author records, all these records are freely discoverable and available via OAI-PMH [16] in MarcXML. With the Karslruhe Institute of Technology (KIT) we know at least one larger non-JOIN²-partner that reuses those records for its own systems.

Technically, JOIN² is based on INVENIO [35], a software developped for large scale repositories by CERN. JOIN², however, was customized to add some features not found in INVENIO. A flavour of Marc21 (expressed in XML) is used as internal data format allowing for very rich metadata. Additionally, while every project partner operates its local installation, the project established a common code base and regular roll out procedures to keep all parterns up to date [36]. This also allows for developments, once implemented, to be used by all.

2.2 Publication Costs: Symbiosis of JOIN² and openCost

JOIN² relies on a jointly developed and constantly reviewed, as well as expanded, infrastructure. Together we work on codes, rollouts, configurations, and share our individual expertise. The following section will depict which implications and workflows regarding data acquisition, provision and internal evaluation are to be considered to organize the publication costs mangement and monitoring of JOIN². The development JOIN² has experienced till today's handling will be illustrated, accompanied by a description of what openCost [37] implies for JOIN² and what the prospects are for strategic further development in the future.

Pretty early on, JOIN² realized that the handling of publication charges, both Open Access as well as Closed Access, are a central service that libraries have to provide [38]. Especially the research centres in JOIN² have a long tradition in this regard stemming from the reprint-era, where reprints, like other literature material, were centrally acquired and managed by the libraries via a dedicated budget. Those budgets usually form the core of the publication funds established decades later.

With the rise of APC funded Open Access it was further realized that a centralized management and monitoring is required, and that a repository, especially if it offers a rich data model and high quality metadata like JOIN², is a natural place to manage these budgets [39–41]:

- In JOIN² the repository is usually *also* used for evaluations, hence scientists have to report their publications and a major goal is *completeness* with regard to the yearly scientific output (see also [39]).

  Bluntly put: If an institution spends an amount of money as for a publication to appear, and then fails to report this very publication as its own achievement, the workflow needs improvement. Hence it makes sense to handle charges centrally at the point that is also responsible for the bibliography and think both requirements as two sides of the same coin [39].

- Some institutions require a dedicated approval process *before* a publication is handed in. Starting this already in the repository/publications database again helps to ensure completeness and avoids the necessity to key in the data multiple times. As such an approval would require at least the preprint this also offers a good chance for Green Open Access later on.

- If a publication is bought by the library and hence is available under a free licence it should be available at the institution via the repository as well.

  - From a user's perspective one should find the article attached without any further forwarding to some publisher site to search there again for the download option.
  - Full texts further the visibility of the repository and hence the institution in search engines.
  - The institution should always have an archival copy.
  - The fine grained metadata available due to other requirements (e. g. reporting) is useful for the monitoring of costs.

- The usual repository infrastructure, like the OAI-PMH [16] interface, is readily available and can be used to make payment data available and visible and thus further transparency. This also ensures a seamless and fast possibility to share and reuse data.

- To store cost data along with publication metadata ensures consistent allocation of costs at the origin using the invoice as source.

  Our approach thus differs from the proposal of e. g. subsection 6.6.1 and avoids issues with assigning sums to items later on as outlined in e. g. subsection 6.1.2.

To maximise the insights from initiatives like OpenAPC [15] or the Open Access Monitor [42] however, the association of costs to the item are mandatory.

- As the cost data is stored together with the original article at the institution that handled (at least one of) the invoice(s), it is clear from the start who paid what. This avoids the difficulty of identifying the paying institution from the beginning. (For those challenges see section 5.2).

  One should note however, that models like SCOAP³ [6] still pose problems in this association as there is no single institution paying.[7]

Similar conclusions as in JOIN² were drawn independently at the University Library of Regensburg at about the same time (see also chapter 1 for the current status there). Discussions with OpenAPC as a natural user for exposed cost data finally led to a proposed infrastructure [43, 44] already in 2017 that will finally be realized in openCost [37, 45].

At JOIN² a simple module to add cost data to publications was implemented in about 2016 [40]. This was initially led by Forschungszentrum Jülich, but quickly adopted by DESY and other partners later on. Also triggered through requirements of the Helmholtz Association, represented by JOIN²-partners involved, as well as to replace and satisfy further local solutions [46].

JOIN² uses a very pragmatic approach, much in line with the findings of section 7.5. We try to add

- a minimal amount of additional data (namely cost-related values)
- at the point of their creation (namely the invoice)
- in a step that needs to be done anyway (namely, the clearing of the invoice by the library)
- *before* this invoice goes to central accounting.

Thus, within the JOIN²-approach cost data is collected at the time they are incurred, using the invoice as the source and not, for example, aggregations from the accounting system or further finance divisions of the individual institute. In this sense it differs from AT2OA2's description in subsection 6.6.1, relying on SAP as resource.

[7] In the context of monitoring the DFG program "Open Access Publikationskosten" [1, 14] this becomes quite apparent.

It is worth noting that the so called APC Module [40] of JOIN² always handled data well beyond Gold OA fees. Almost all normalized inputs are configurable to enable the system to handle whatever we come across, as we always wanted to cover all costs paid [38]. Hence it always contained categories like color charges hybrid OA, cover carges, costs for reprints, permissions (licencing fees) or page charges.

For monitoring JOIN² relies on metadata that were already added to the repository to describe the bibliographic item. Due to the high amount of automatic normalization in the submission process and the use of the repository also for research assessment, evaluations can be done on detailed aspects without the need to extract data from systems that have nothing to do with bibliographic information or where sums first need to be split to individual items by complex and error prone logics.

On the technical side all publication charges are handled like a book purchase. That is, owing to the Marc21 internal format, JOIN² adds pretty much the same Marc holdings block as it is done in a library system. The sole difference is, that the item is defined as the bibliographic entity just described, while each holdings block refers to a part of the individual payment. This allows to hold partial cots as shown on the invoice while a traditional holdings entry would only contain the overall sum. Figure 2.2 shows a cost block for a publication with four identifiable cost categories. Note that the Colour charges are in the same order of magnitude as the (hybrid) publication fee.

All other necessary metadata (see e. g. section 1.6, or [12]) are already stored during the submission to the repository and are thus readily available. The definitions for the necessary cost elements were done in close collaboration with those colleagues who actually handle the invoices. In other words *"What do you see on the bill?"*[8] is the key guiding question still frequently asked when adoptions for changes are required (e. g. transformative agreements).

In addition, for those colleagues involved in the collection and entering of cost data into the repository, the Electronic Journals Library [47] serves as a substantive frame of reference. The data in EZB is acquired by a cooperative effort (see section 9.3) and builds one corner stone of JOIN²'s journal identification system. As additionaly values are required, JOIN² explicitly supports the functional expansion of the EZB, intended within the framework of openCost in order to also provide information on the publication status in the future.

[8] now expanded by *"What does the dashboard show you?"*

```
876 7_  |c 1194.99
        |v 11.16
        |e Colour charges
        |d 2022-04-22
        |j Zahlung erfolgt
        |p 448457
        |9 2022-04-08
        |z additional figures
876 7 _ |c 32.16
        |v 0.30
        |e Other
        |d 2022-04-22
        |j Zahlung erfolgt
        |p 448457
        |9 2022-04-08
        |z payment processing fee
```

```
876 7_ |c 2050.15
       |v 19.14
       |e Hybrid-OA
       |d 2022-04-22
       |j Zahlung erfolgt
       |p 448457
       |9 2022-04-08
       |z CCBY license
876 7_ |c 597.49
       |v 5.58
       |e Colour charges
       |d 2022-04-22
       |j Zahlung erfolgt
       |p 448457
       |9 2022-04-08
       |z first figure
```

Figure 2.2: This cost block (slightly shortened internal format) shows four parts of a single invoice. As indicated by the fee type Hybrid-OA the article in question was published in a (at time of payment) hybrid journal. Besides the Hybrid-OA fee, the publisher also billed Colour charges twice: For the first figure 597.49 EUR +VAT and another 1194.99 EUR +VAT for all subsequent figures. Additionally, 32.16 EUR were charged for payment processing (payment processing fee). The orignal payments (not visible here) were done in USD and only later converted by means of the central accounting to EUR. (Note: The complete cost block also features some internal values, like the link to central accounting, budgets charged etc..)

In a first step journal categories as outlined in subsection 9.5.2 will allow JOIN²'s journal records to know the category of a journal in question (e. g. to identify mirror journals). This streamlines fee processing as certain budgets are only availabe for some given categories.

In the transition phase towards the *Open Access default* transformative agreements play a special role. As these agreements usually consist of a large number of ever changing individual journals. A simple interface that can be queried using journal identifiers (ISSN-, ZDB- or EZB-IDs are used extensively in JOIN²) and returns if the journal in question is subject to some contract and whether one's institution has subscribed to that contract has been a long standing desideratum, that basically exists at least since the very first DEAL agreement with Wiley. Hence, the ideas outlined in subsection 9.5.3 will ease up things for the staff considerably.

Even in a world where transformative agreements are a thing of the past, a service that returns information on specific agreements (currently misnamed 'memberships') will be a standing requirement and again the ideas of subsection 9.5.3 could come to the rescue.

At the moment this information is kept in internal lists which is cumbersome and error prone compared to a community created and curated database. Once the EZB eventually adds a Web-API providing this information upon request by some ID, changes will show up seamlessly within the JOIN²-systems. Combined with the glossary that EZB is about to build up, the currently ambigious definitions can be sorted out and repository staff will have a reliable source of information and a central point of reference for identifying the journals status at hand. This is further improved by the central data handling envisioned.

For the provision of JOIN²'s publication costs data an XML exchange format is utilized, which is to be adapted continuously to the current openCost version. An intermediate version can be seen in detail in Figure 2.3. The cost schema, that is based on a controlled vocabulary, is structurally capable of covering multiple fee types as well as multiple fee entries, different currencies and several identifiers. Always with the central focus on automatic data exchange, the repositories allow the provision of cost data in a machine-readable form for harvesting via the OAI-PMH interface [16, 48].

Besides the continuous implementation of the future openCost standards, JOIN² is facing further steps to be realized. Taking up the headings of harvesting and automated data exchange, all JOIN²-partners will be able to deliver their data to OpenAPC or provide it to all other interested parties in the future. The current process of manually creating tables and sending them in by mail did not work

```
<oc:publication>
 <oc:amounts_paid>
  <oc:item>
   <oc:amount_invoice currency="USD">4215.0</oc:amount_invoice>
   <oc:amounts_paid>
    <oc:amount_paid currency="EUR" type="hybrid-oa">2050.15</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="vat">19.14</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="colour charges">597.49</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="vat">5.58</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="colour charges">1194.99</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="vat">11.16</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="other">32.16</oc:amount_paid>
    <oc:amount_paid currency="EUR" type="vat">0.30</oc:amount_paid>
   </oc:amounts_paid>
   <oc:date_paid>2022-04-22</oc:date_paid>
   <oc:publisher>Soc.</oc:publisher>
   <oc:invoice_number>448457</oc:invoice_number>
  </oc:item>
 </oc:amounts_paid>
 <oc:oa_status>hybrid</oc:oa_status>
 <oc:primary_identifier>
  <oc:doi>10.1364/OL.448457</oc:doi>
 </oc:primary_identifier>
 <oc:secondary_identifiers>
  <oc:id type="local">PUBDB-2021-04978</oc:id>
  <oc:id type="oai">oai:bib-pubdb1.desy.de:472306</oc:id>
  <oc:id type="doi">10.1364/OL.448457</oc:id>
  <oc:id type="doi">10.3204/PUBDB-2021-04978</oc:id>
 </oc:secondary_identifiers>
 <oc:type>journal article</oc:type>
 <oc:institutions>
  <oc:institution type="short_name">desy</oc:institution>
  <oc:institution type="ror">https://ror.org/01js2sh04</oc:institution>
 </oc:institutions>
</oc:publication>
```

Figure 2.3: Sample record of Figure 2.2 expressed in an early version of openCost XML.

well due to limited resources.

The management in a JOIN² repository furthermore enables local Open Access publishing, if possible (Green Road). With cost transparency as one of the central goals of the Open Access transformation, the described efforts are now gaining importance beyond one's own institution with reporting for the DFG Open Access Publication Costs program [14]. Meeting the funding requirements thus often becomes a first intermediate goal on the way to openCost. It is worth noting that the current layout of cost handling in JOIN² already fulfills all requirements of this program and enabled an almost automatic creation of the monitoring table already for the first round of this program. In addition to certain challenges, this program also offers some project partners the first opportunity to establish the structural adjustments they have long been striving for. In this respect, the JOIN²-partners are also seeking to further enhance their monitoring.

Given the diverse partners JOIN² will provide an ideal testing ground within openCost and serve as a reference frame. Regular user meetings generate, scrutinize and review the results generated in openCost and subject them to practicality. In this way, openCost benefits from the discourse of the participating institutions, which test the initial project results in their own practical world and exchange experiences. The JOIN² repositories, along with those of the university libraries of Regensburg and Bielefeld, are thus the first to actively implement the openCost standards.

Moreover the automized reporting, including non Open Access fees, should be improved. How to optimize workflows and advance centralization will be an ongoing part of the debate. In this regard, the developments of the Transform2Open team (see e. g. [49] and chapter 8) in particular will be expected with great interest. With openCost heading for the more technical part and Transform2Open diving into workflows both projects are complementary and JOIN² expects quite a few synergies.

2.3 Organizational Challenges & Library's Engagement

2.3.1 Central Invoicing

A common goal in JOIN² is cost transparency within and between institutions that should be improved to enable comparisons and raise awareness of publication costs in general as well as the various components these charges are consisting of. To promote this, central invoicing is required.

It is thus a joint objective that the library assumes responsibility for publishing costs and centrally manages the entire process of publication charge handling. It should form a one-stop-shop for the researchers on campus. It is a common misunderstanding that this requires one single central fund to be the only source for payments. In fact, various budgets may be eligible for a single publication, so our view is in line with section 1.5 (see also the figure in [39], p. 17 for a very early view).

For example at DESY the library currently pays all Open Access charges via its own budget. But the library of course also handles publication charges related to Closed Access (e. g. colour charges, cover charges) or hybrid Open Access (also beyond so called transformative agreements). However, those costs are not covered by the library budget, but are charged to the budgets of the originating groups on campus. In these cases the library is a service facility for the processing. Once the invoice is checked by the library the costs are charged by the library to the cost centres accordingly, which also allows for inner-institutional cost splitting, e. g. depending on the cost type or the availability of third-party funding (e. g. in case of EU funded research). That is, while all invoices are handled centrally by the library, the actual budgets charged may belong to individual institutes on campus, central funds or third-party funding or a combination of all of those. In view of the current German DFG project Open Access Publication Funding [14] one can easily envision some amount to be charged to the publication fund, some amount to the program's cost centre and some non-OA-related charges (e. g. cover charges) to an institute's budget.

2.3.2 The repository

The JOIN² partner institutions [22] are working collaboratively on their approach to manage publication costs and there is a common understanding to use the repositories to this end for the reasons outlined in section 2.2. However, the precise organization for the handling of costs and internal workflows differs sometimes considerably, e. g. with regards to:

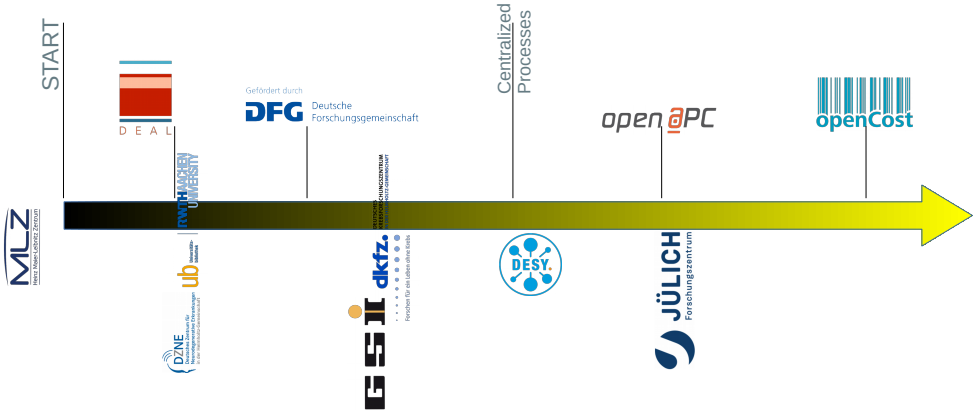

Figure 2.4: The road to openCost: In JOIN² the process to implement publication charge handling is at quite different stages. While Forschungszentrum Jülich is already rather advanced, delivering data (manually) to OpenAPC for some time now and being responsible for the DFG monitoring [1], other institutions are about to build up processes just now. At GSI and DFKZ the DFG Program Open Access Publication Funding [50] is a starting point to build up those structures. A few already hold cost data for years. Sometimes, however, those are not reflecting all publications due to decentralized processes.

- current status of publication costs handling on site (see also Figure 2.4)
- data acquisition, or "Who is actually handling the invoice?"

  To gain the most of the integrated workflow, it is common to handle it in the group responsible for scientific publishing, but sometimes it is done by the acquisition department.
- monitoring (similar to data acquisition)
- funding (central, decentral, dedicated funds, …)
- funding criteria (e. g. corresponding authors)
- number of personnel (the "group" may consists of only half a position)

Even though the outset in each library is quite different the experiences over the past years also show that there are quite some similarities. They may just not

be visible at first glance. It was also observed that some procedures that initially were handled differently converged over time to some best practise approach. For example, even though almost all parameters of the mentioned APC module (see section 2.2) are configurable and initially showed several different readings, over time they basically converged to the same set of values.

As many of the Helmholtz libraries are comparatively small while at the same time their institutions are very productive in research output, streamlining the overall process to improve efficiency is always a guiding principle and good enough may be better than perfect but unmanageable (see also section 7.5). Getting the library on board early on has several advantages in this regard, even for the library:

- It gives the opportunity to assist with author rights, copyright and licensing questions or implement measures against predatory publishing, e. g. as an outlet for results from SP4 of subsection 6.1.4.
- It offers the possibility to avoid unnecessary costs, e. g. by pointing at a Green OA option instead of expensive Hybrid OA to fulfill funders' requirements.
- The library can not only provide clarification for authors, but also optimizes and ensures compliance with or feasibility of the guidelines presupposed by the own institution or funders.
- It allows to inform authors on campus seamlessly about the payment (status, amount etc.).
- Usually, it ensures that the invoices are correct in the first place and the necessity to ask for corrections is reduced (e. g. VAT handling, billing address, special regulations due to a contract the authors are unaware of while the library is managing them).
- The library staff just processes *way more* invoices by many more different publisher than even the most productive groups on campus. Therefore, they usually know all glitches within the interfaces involved as well as special regulations due to contracts.
- In view of the Helmholtz Open Science Policies the institutions have to report on their Open Access rates with a 100% goal to be reached in 2025 [51, 52]. As this includes Green OA it is sensible to derive the reports from the repository.

2.3.3 Information Budget

In view of the information budget (see also section 1.4 and in more details chapter 4) the actual approach in JOIN² can be considered consistent with the core of Figure 4.1 taking section 4.2.2.1 into account. Therefore, JOIN² currently lives on the basic costs as defined in section 7.3. It does not deal with the outer shells and may never do, as the repository is focused on single publications. On the other hand the repository approach offers detailed, fine granular data for an important and most likely growing part of the overall information budget. Additionally, the mentioned central invoicing (subsection 2.3.1) allows to unearth expenditures for information that went by unseen in the past. A quick look at individual expenditures for e. g. Closed Access publishing already demolishes the myth that only Open Access requires publication charges.

2.4 Conclusion

With the ambition to actively contribute to the transformation process of scholarly publishing in accordance with FAIR[9] principles JOIN² decided to identify collective strategies to organize their publication costs management centralized in their entirety. The JOIN² partner institutions can reflect on many years of successful cooperation and will continue to work on optimizations as well as the further realization of the openCost. Although they formally share the same baseline, that means the context of their repository infrastructure, they nevertheless deal with rather different institution-specific requirements: The publication costs scenario varies depending on the institutional conditions, e. g. data acquisition, the interaction between different departments involved in cost processing, funding and funding criteria, number of personnel, and constraints on or urgency of action.

Thus, even though the JOIN² institutions must be fetched from different stations along their road to publication costs transparency (Figure 2.4), the cooperation could be seen as *proof of concept* for realizing openCost goals. Even though individual preconditions are given, the formulated problems and challenges are mostly the same. Obstacles can be overcome and replaced by standardized solutions. openCost functions as a self-contained, independent module that is integrated as a building block into the JOIN² infrastructure, and which will be continuously adapted in the future in line with the newest joint developments. Using JOIN² as an example, various advantages could be demonstrated speaking in favor for handling

[9] **F**indable, **A**ccessible, **I**nteroperable, **R**eusable

the management and monitoring via the library in the institutional repository, ranging from data acquisition to internal evaluation and data provision.

References

[1] Forschungszentrum Jülich. *Monitoring DFG-Förderprogramm „Open Access Publikationskosten". Erläuterungen zum Datenschema*. Apr. 6, 2023. URL: https://www.fz-juelich.de/de/zb/downloads/open-science/monitoring-oa-publikationskosten-datenschema-erlaeuterungen-2023-01-25.pdf/@@download/file (cit. on pp. 2, 16, 23).

[6] CERN. *Sponsoring Consortium for Open Access Publishing in Particle Physics. SCOAP³*. Apr. 6, 2023. URL: https://scoap3.org (cit. on pp. 3, 16).

[12] Heinz Pampel. *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten*. de. Arbeitspapier. 2019. DOI: 10.2312/OS.HELMHOLTZ.006 (cit. on pp. 6, 17, 47, 72, 75).

[14] Deutsche Forschungsgemeinschaft. *Funding Programme: "Open Access Publication Funding"*. URL: https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/open_access_publikationskosten/ (visited on 2022-12-13) (cit. on pp. 8, 16, 21, 22, 47, 49).

[15] Bielefeld University Library and DINI working group Electronic Publishing. *OpenAPC*. en. URL: https://openapc.net/ (visited on 2022-12-06) (cit. on pp. 8, 16, 103).

[16] Open Archives Initiative. *Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14*. Jan. 8, 2015. URL: http://www.openarchives.org/OAI/openarchivesprotocol.html (visited on 2023-04-06) (cit. on pp. 8, 14, 15, 19).

[19] JOIN². *JOIN² - Just anOther INvenio INstance To Join*. en. URL: https://join2.de/ (visited on 2022-12-08) (cit. on pp. 11, 12, 38, 103).

[20] A. Wagner. "Veröffentlichungsdatenbank und Volltextrepositorium". In: *BIT online* 14 (2011), pp. 45–48 (cit. on p. 12).

[21] A. Wagner. “Ein neues JUWEL?: Publikationsmanagement für Wissenschaft und Administration”. de. In: 33. Arbeits- und Fortbildungstagung der Arbeitsgemeinschaft der Spezialbibliotheken e.V. - Sektion 5 im Deutschen Bibliotheksverband, Jülich. 2011 (cit. on p. 12).

[22] JOIN². *JOIN² Instanzen*. 2022. URL: https://join2.de/Main/Instanzen (visited on 2022-12-06) (cit. on pp. 12, 22, 38).

[23] Alexander Wagner and Robert Thiele. “Invenio@HGF – Collaborative repository infrastructure”. In: Open Repositories 2014, Helsinki (Finland), 8 Jun 2014 - 13 Jun 2014. June 8, 2014. DOI: 10.3204/DESY-2014-02793 (cit. on p. 12).

[24] Alexander Wagner. “JOIN² - What it is and what it does”. In: 2016. DOI: 10.3204/PUBDB-2016-01331 (cit. on p. 12).

[25] Alexander Wagner. “Publication database”. In: EUCALL Workshop: User Access at Advanced Laser Light Sources, Dolní Břežany (Czechia), 21 Sep 2017 - 22 Sep 2017. Sept. 21, 2017. DOI: 10.3204/PUBDB-2017-10906 (cit. on p. 12).

[26] Laura Baracci, Alexander Wagner, and JOIN2. “JOIN² : A Publication Database and Repository based on Invenio”. In: XXVII International Symposium on Nuclear Electronics & Computing, Budva Becici (Montenegro), 30 Sep 2019 - 4 Oct 2019. Sept. 30, 2019, pp. 51–57. DOI: 10.3204/PUBDB-2019-04497 (cit. on p. 12).

[27] Irina Filozova et al. “JINR Open Access Repository based on the JOIN² Platform”. In: *Supplementary Proceedings of the XXII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020)*. Vol. 2790. Data Analytics and Management in Data Intensive Domains 2020, Voronezh (Russia), 13 Oct 2020 - 16 Oct 2020. Aachen, Germany: RWTH Aachen, Oct. 13, 2020, pp. 142–155. DOI: 10.3204/PUBDB-2021-00818 (cit. on p. 12).

[28] Alexander Wagner et al. “JOIN² empowering a bibliographic infrastructure to support scientists”. In: CERN Workshop on Innovations in Scholary Communication, Geneva (Switzerland), 17 Jun 2015 - 19 Jun 2015. June 17, 2015. DOI: 10.3204/PUBDB-2015-02188 (cit. on p. 12).

[29] Alexander Wagner. “JOIN² – A scientists toolbox”. In: Dubna (Russian Federation), 13 Jan 2018 - 21 Jan 2018. Jan. 13, 2018. DOI: 10.3204/PUBDB-2018-00618 (cit. on p. 12).

[30] Alexander Wagner. “JuSER - Autorenhandling”. In: DINI-/Helmholtz-Workshop „Autorenidentifikation am Beispiel von ORCID“, , 15 Oct 2012. Oct. 15, 2012. DOI: 10.3204/PUBDB-2017-00856 (cit. on pp. 12, 13).

[31] Alexander Wagner. “Authority Control in Invenio”. In: 2nd Invenio User Group Workshop, Jülich (Germany), 18 Nov 2013 - 20 Nov 2013. Nov. 18, 2013. DOI: 10.3204/PUBDB-2017-00870 (cit. on pp. 12, 13).

[32] Rebecca Mehl et al. *Repositorium, Publikationsdatenbank und Katalog: Integrierte Lösungen für kleine Bibliotheken*. Tech. rep. 2018, p. 6. DOI: 10.3204/PUBDB-2019-01442 (cit. on p. 12).

[33] National Library of Medicine, ed. *Pubmed Central*. URL: https://www.ncbi.nlm.nih.gov/pmc (cit. on p. 13).

[34] Cornell University, ed. *arXiv*. URL: https://arxiv.org (cit. on p. 13).

[35] CERN. *Invenio*. URL: https://invenio-software.org/about/ (cit. on p. 14).

[36] Alexander Wagner. “Invenio @ HGF - Technical background”. In: Invenio Developers Forum, Geneva (Switzerland), 16 Jun 2013. June 16, 2013. DOI: 10.3204/PUBDB-2017-00861 (cit. on p. 14).

[37] openCost. *openCost: Projekt*. URL: https://www.opencost.de/projekt/ (visited on 2022-11-27) (cit. on pp. 14, 16, 112).

[38] Alexander Wagner. “Publikationsfonds”. In: Open Access Tage 2011, Regensburg (Germany), 5 Oct 2011 - 5 Oct 2011. Oct. 5, 2011. DOI: 10.3204/PUBDB-2017-00849 (cit. on pp. 14, 17).

[39] Alexander Wagner. “Publizieren ist nicht genug”. In: vol. 22. Schriften des Forschungszentrums Jülich. Reihe Bibliothek / library. 7. Konferenz der Zentralbibliothek des Forschungszentrums Jülich, Jülich (Germany), 14 Jun 2016 - 16 Jun 2016. Jülich: Zentralbibliothek Forschungszentrum Jülich, June 14, 2016, pp. 23–47. DOI: 10.3204/PUBDB-2016-01819 (cit. on pp. 14, 15, 22).

[40] Alexander Wagner. “Article Processing Charges and OpenAPC”. In: 4th Invenio User Group Workshop, Garching (Germany), 21 Mar 2017 - 24 Mar 2017. Mar. 21, 2017. DOI: 10.3204/PUBDB-2017-01358 (cit. on pp. 14, 16, 17).

[41] Alexander Wagner et al. “APC-Verwaltung im institutionellen Repositorium”. de. In: 107. Deutscher Bibliothekartag - offen & vernetzt, Berlin (Germany), 12 Jun 2018 - 15 Jun 2018. June 12, 2018. DOI: 10.3204/PUBDB-2018-02211 (cit. on p. 14).

[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, 112).

[43] Gernot Deinzer et al. “Including Open Access Information in institutional repositories”. In: CERN Workshop on Innovations in Scholary Communication, Geneva (Switzerland), 21 Jun 2017 - 23 Jun 2017. June 21, 2017. DOI: 10.3204/PUBDB-2017-05961 (cit. on pp. 16, 37).

[44] Gernot Deinzer et al. “Including Open Access Information In Institutional Repositories”. In: Open Access Tage 2017, Dresden (Germany), 11 Sep 2017 - 13 Sep 2017. Sept. 11, 2017. DOI: 10.5281/zenodo.887318 (cit. on p. 16).

[45] openCost. *openCost*. URL: https://www.opencost.de (visited on 2022-12-13) (cit. on pp. 16, 37, 46).

[46] Dagmar Sitek, Alexander Wagner, and Corinna Mescherowsky. “Open-Access-Publikationskosten @ JOIN²”. In: Open-Access-Tage 2022, Bern (Switzerland), 19 Sep 2022 - 21 Sep 2022. Sept. 19, 2022 (cit. on p. 16).

[47] Universitätsbibliothek Regensburg. *Elektronische Zeitschriftenbibliothek (EZB)*. de. URL: https://ezb.ur.de/ (visited on 2022-12-06) (cit. on pp. 17, 102).

[48] Alexander Wagner. “APC-Verwaltung im institutionellen Repositorium”. de. In: *GMS Medizin, Bibliothek, Information* 18.3 (Sept. 10, 2018), Doc21. DOI: 10.3205/mbi000422 (cit. on p. 19).

[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 (cit. on pp. 21, 46, 95, 98, 121).

[50] Deutsche Forschungsgemeinschaft. *Open Access Publication Funding: Guidelines and Supplementary Instructions*. DFG form 12.21. 2022 (cit. on pp. 23, 112, 115).

[51] *Open Access Policy of the Helmholtz Association. Version 1.0. Approved in the 119th General Assembly of the Helmholtz Association on 20-21 September 2022*. 2016. 2 pp. DOI: 10.48440/os.helmholtz.018 (cit. on p. 24).

[52] *Helmholtz Open Science Policy. Version 1.0. Approved in the 119th General Assembly of the Helmholtz Association on 20-21 September 2022*. 2022. 9 pp. DOI: 10.48440/os.helmholtz.056 (cit. on p. 24).

3

Extension of the OpenAPC infrastructure as part of the openCost project

Julia Bartlewski[1]
Bielefeld University Library

Christoph Broschinski[2]
Bielefeld University Library

DOI: 10.3204/DESY-PROC-2023-01/03

Abstract

The OpenAPC initiative collects and disseminates datasets on fees paid for open access publishing under an open database license. OpenAPC is operated by Bielefeld University Library. The aims of OpenAPC are transparency and reproducibility of OA costs, as well as to illustrate the development of costs over time.

The current OpenAPC metadata schema is to be extended as part of the openCost project in order to enable collecting additional cost information, such as color orpage charges, in a structured and standardized way. So far, data is mainly reported as csv-files to OpenAPC, either via email or as pull requests on GitHub. As an important element of the project, OpenAPC plans to expand the harvesting of metadata and costdata. In the openCost project, this will be done exemplarily for the repositories of the partner

[1] https://orcid.org/0000-0001-5959-4999
[2] https://orcid.org/0000-0003-1972-7587

institutions. The talk presents the first results and serves as a starting point for the extension of the metadata schema.

3.1 Introduction

OpenAPC is an open data project, established at Bielefeld University Library in 2014 [53]. It focuses on the aggregation of cost data on open access (OA) publishing under an Open Database License [54]. In the initial phase, the project concentrated on aggregating Article Processing Charges (APCs) from German universities that operated publishing funds supported by the Open Access Publication Program of the German Research Foundation (DFG). From 2015 until 2018 OpenAPC was part of the project Transparent infrastructure for open access publication fees (INTACT) which was funded by the DFG [55] As of October 2020, OpenAPC has been established as a long-term service at Bielefeld University Library.

Publication Type	Count	Aggregated Sum	Contributing Institutions
Articles	167 505	322 494 631 €	381
Monographs	1 466	9 476 513 €	27

Table 3.1: OpenAPC cost data as of 04/10/2022

The project started with aggregating APCs [56] and two years later, in 2016, OpenAPC begun to collect and disseminate data on articles published under transformative agreements [57]. Since 2020, OpenAPC also aggregates Book Processing Charges (BPCs) [58] For these three different types OpenAPC maintains three different data sets. Table 3.1 shows some facts and numbers regarding the APC and BPC data sets.

3.2 OpenAPC: Aims and principles

The overall aims of OpenAPC are to create cost transparency and comparability between institutions as well as to enable transparent and reproducible reporting for institutions and funders. Furthermore, OpenAPC also tracks cost development over time.

institution	period	euro	doi	is_hybrid	publisher	journal_full_title	issn	issn_p	issn_e	issn_l	license_ref	in_crossref	pmid	pmcid	ut	url	doaj
Bamberg	2013	1372	10.2147...	FALSE	Dove Medical...	Journal of...	1178-7090	NA	1178-7090	1178-7090	http://creative...	TRUE	2378...	PMC3...	ut:00020...	NA	TRUE
Bamberg	2013	960	10.3389...	FALSE	Frontiers...	Frontiers in...	1664-1078	NA	1664-1078	1664-1078	NA	TRUE	2376...	PMC3...	ut:00033...	NA	TRUE
Bamberg	2013	960	10.3389...	FALSE	Frontiers...	Frontiers in...	1664-1078	NA	1664-1078	1664-1078	NA	TRUE	2384...	PMC3...	ut:00033...	NA	TRUE

Reported by institution | Crossref | Europe Pubmed Central | ISSN Organisation | Web of Science | DOAJ

Table 3.2: OpenAPC metadata schema for journal articles

With its service, OpenAPC intends to provide data as a foundation for policy making, since taking decisions within the Open Access transition framework requires reliable and resilient figures and facts.

OpenAPC takes on an open science based approach, which pursues our main principle of being as transparent as possible in everything we do. The *Open* in OpenAPC refers not only to the data itself, but also to all the code and scripts we write, the workflows we use, and the history of the project itself. This is accomplished by using the version control system git. All files which are relevant to the project are stored and revised in a repository on the platform GitHub [59], so one can basically trace all the code back to its origins in 2014, when the first commit was made [60]. GitHub serves as the main data storage for all the data received. Due to the release under an Open Database License, all collected data can be reused by anyone, e. g. to perform one's own analyses and calculations.

Another principle of OpenAPC is the use of many automated processes. On the one hand, this is due to the high volume of ingested data, which can thus be processed more efficiently, and on the other hand, OpenAPC wants to apply a low entry barrier for new participants.

3.3 OpenAPC: Data schema and workflows

Since all collected data has been voluntarily provided by external participants, OpenAPC puts a lot of emphasis on making participation and data submission as easy as possible. There is no formal admission or membership of institutions, submitting data in a simple table format (CSV) and naming a contact person is sufficient. There are data submissions from individual universities or institutes, scientific organizations, research funders or consortia [61].

Table 3.2 shows the current OpenAPC metadata schema for journal articles, which illustrates the high level of automation that is applied.

The metadata schema consists of 18 fields, but only the first five ones are required to be reported by participating institutions [62].

- institution top-level organisation which covered the reported costs
- period year of payment
- euro the amount that was paid in EURO
- doi article DOI
- is_hybrid a Boolean indicator if the journal is Hybrid or Gold OA

All remaining data fields are automatically enriched using external sources, notably Crossref, Europe PubMed Central, the DOAJ, the Web of Science, and the ISSN organization [63]. This automated enrichment process has a number of advantages. First of all, the workload for the participating institutions remains manageable, since they only have to provide a minimal data set. Furthermore, this approach improves the consistency of publisher names and journal titles, which is crucial for further analysis and visualization. Finally, during the enrichment process, the submitted data is normalized and reformatted in such a way that the results are compliant with the OpenAPC data schema. The enriched files are also stored in the official project repository on GitHub. The contents of all these enriched files are aggregated into one main CSV file, which forms the OpenAPC data set. The whole OpenAPC data set is automatically tested for errors whenever new articles are integrated, including identifier checks (DOIs/ISSNs, via regular expressions), logical checks (Journals listed in the DOAJ cannot be hybrid) and searching for DOI duplicates. Serious errors (like DOI duplicates with other institutions) are reported back to the participants.

The submission of data on OA monographs works in the same way. The BPC data set is the newest addition to OpenAPC and collects data on BPCs (Book Processing Charges). The schema consists of 13 fields, with the first five fields being mandatory (see Table 3.3).

Table 3.3: OpenAPC metadata schema for books

institution	period	euro	doi	is_hybrid	publisher	journal_full_title	issn	issn_print	issn_electronic	issn_l	license_ref	indexed_in_crossref	pmid	pmcid	ut	url	doaj	agreement
Aberystwyth U	2015	NA	10.1007	TRUE	Springer Nature	Soft Computing	1432-7643	1432-7643	1433-7479	1432-7643	http://creativ	TRUE	NA	NA	ut:0	NA	FALSE	Springer Compact
Aberystwyth U	2016	NA	10.1007	TRUE	Springer Nature	Geoheritage	1867-2477	1867-2477	1867-2485	1867-2477	http://creativ	TRUE	NA	NA	ut:0	NA	FALSE	Springer Compact
Aberystwyth U	2016	NA	10.1007	TRUE	Springer Nature	Studies in Philosop	0039-3746	0039-3746	1573-191X	0039-3746	NA	TRUE	NA	NA	ut:0	NA	FALSE	Springer Compact
Aberystwyth U	2016	NA	10.1007	TRUE	Springer Nature	Language Policy	1568-4555	1568-4555	1573-1863	1568-4555	http://creativ	TRUE	NA	NA	ut:0	NA	FALSE	Springer Compact

mandatory

Table 3.4: TA metadata schema (OpenAPC)

In both schemas only one field (euro) is designated for recording costs. Since OpenAPC collects and publishes data on expenditures incurred by institutions for OA publishing, defining the term costs was an important step in the project's initial setup [64]. Cost data should be as consistent and comparable as possible, while being easy to collect and report. For consistency, OpenAPC only collects data on fees paid for journal articles (APCs) and books (BPCs). Other publication types such as conference papers or book chapters are not collected. All reported publication fees are gross values, including taxes or discounts. Additional expenses not related to the OA component, like page charges or submission fees, should be excluded.

The Transformative Agreements (TA) data set contains metadata on OA journal articles published under transformative agreements instead of being paid with classical APCs. The TA schema consists of 19 fields and is very similar to the OpenAPC one. It includes an additional column agreement, indicating the name of the contract under which the article was published. The Euro field is not mandatory as cost and payment modalities can differ a lot across transformative agreements. Therefore most records in the data set do not include cost information (see Table 3.4).

For efficiency reasons, only large data submissions provided by funders or consortia are integrated into the TA data set. A notable exception is data regarding the German DEAL agreements with Wiley and Springer Nature. This is reported individually by the participating institutions. OpenAPC developed workflows to report cost data for this specific use case. The costs per article in Wiley and Springer Nature hybrid journals are calculated by the participating institutions [65].

3.4 OpenAPC: Services and data ingestion

3.4.1 Services

In addition to the GitHub repository, OpenAPC also runs other services. For each submission a post is created in the OpenAPC project blog [66], where the submitted data is analyzed in more detail.

Moreover, auto-generated reports are compiled for each APC submission to provide clues to possible errors in the submitted data. Reported APCs are compared to existing data in the OpenAPC collection. A statistical approach is used to identify articles whose costs differ by more than two standard deviations from the average APC of the respective journal. These reports are sent back to institutions upon every data submission.

The visualization of the data is realized by using interactive treemap graphics [67].

The treemaps present the most easy and intuitive way to browse and inspect the OpenAPC data. It is possible to either view the data for single institutions or the whole data set (Figure 3.1). Furthermore, there are also special treemaps, e. g. for DEAL data. Finally, OpenAPC also operates an OLAP server, which serves both as backend for the treemap server and also offers a fast and efficient way of querying the OpenAPC data via an API [68].

3.4.2 Data ingestion

Participants may provide data to OpenAPC in several ways. The most common option is to send the data via e-mail as a CSV file. Another way is to make a pull request on GitHub. Even though in this case a CSV file is provided as well, this has the advantage of allowing participants a more granular control over their institutional directory on GitHub, where they can make edits and correct errors without the need of direct communication with the OpenAPC team. The most advanced way for institutions is by integrating cost data directly into their institutional repositories and letting OpenAPC harvest it via the OAI-PMH protocol. This approach is technically and organizationally more demanding compared to the others previously mentioned, but once established it offers many advantages for the participating institutions. The direction of the data workflow is reversed: OpenAPC is responsible for collecting the data on a regular basis (the current interval is semiannual), so compiling tables and submitting them on a regular basis is no longer necessary. Detected errors only have to be corrected in the

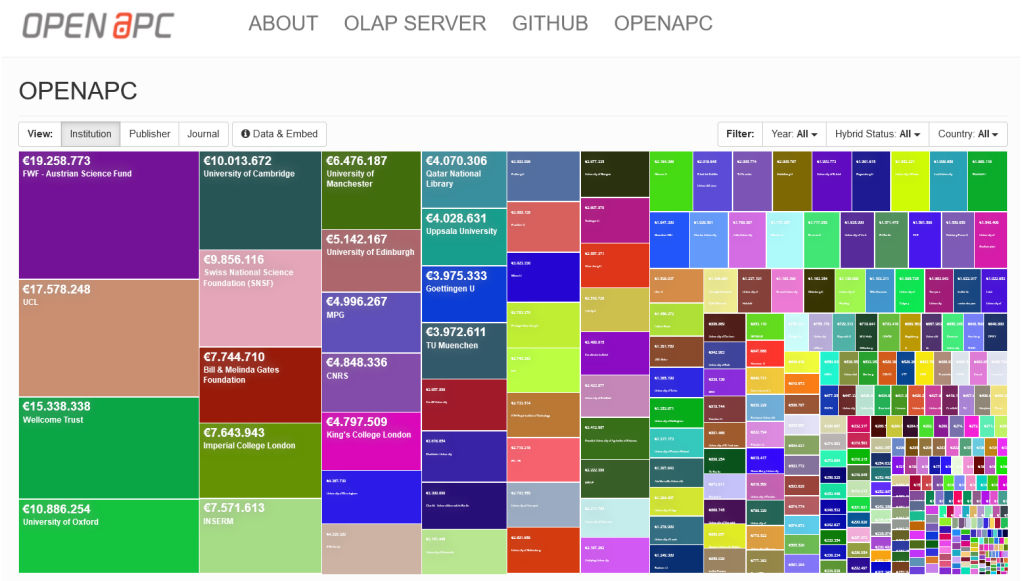

Figure 3.1: Treemap for the OpenAPC data set

institution's own repository, and these changes are automatically applied during the next harvesting. At the moment three institutions provide OpenAPC with data that way (as of October 2022): University of Regensburg, Saarland University and Bielefeld University [43].

However, the XML schema for exchanging cost data is still preliminary and not yet specified. This is one important reason why OpenAPC took part in the project openCost, which is funded by the German Research Foundation (DFG) [45]. One of the aims in the context of the openCost project is to develop a new, universal metadata schema, which can be used for harvesting. This might then serve as a starting point for further institutions to be able to include their cost data directly into their repositories and make OpenAPC harvest it.

So far, in the context of OpenAPC, only the Euro amount as described above has been aggregated, which covers APCs, including taxes and discounts, but without additional costs that are also linked to electronic publishing, such as color or page charges. As part of the openCost project, the OpenAPC infrastructure is to be extended so that these additional costs can also be aggregated in the future [69]. This will contribute to increasing cost transparency within the scope of scientific publishing. Thus, collecting and analyzing these types of costs is intended to show how relevant they are in relation to classical APCs and what share the various

cost items make up in the bigger picture. Therefore, we set up a preliminary pilot project for OpenAPC, in which first test data with additional costs have been integrated.

3.5 Pilot Project: Integrating additional cost data into OpenAPC

For this pilot project, we were able to benefit from the activities of the openCost project partner Deutsches Elektronen-Synchrotron DESY in Hamburg. DESY is part of the JOIN² consortium, a collaborative repository solution currently consisting of eight institutions [19].

JOIN² is based on the software INVENIO. Some participants, including DESY, had already recorded additional costs, such as color charges or page charges, in addition to APCs. Furthermore, the various JOIN² instances already offer OAI-PMH interfaces that allow this data to be harvested [22]. The data we used originates from three JOIN² institutions: Along with data from DESY, we also harvested data from Forschungszentrum Jülich (FZJ)and GSI Helmholtzzentrum für Schwerionenforschung (GSI) We then aggregated this data in our OpenAPC repository and analyzed it, the specific process is described in the following section.

3.5.1 Approach

First, the raw data was harvested via OAI-PMH from the three repositories. Since the data also included closed access publications and non-journal articles, we had to filter out these records beforehand. Once only OA publications remained, we performed the standard OpenAPC metadata enrichment. One crucial question was how to store the additional data fields without impacting the existing OpenAPC dataset. For the purpose of the pilot project we decided to create a special table to store those additional costs and link it to the OpenAPC main data set using the DOI as primary key, this is possible since all OpenAPC sets are duplicate-free in respect to DOIs. All additional cost types which were detected in the JOIN² data were also included into this table.

Table 3.5 shows an excerpt of the DESY data to illustrate this approach. At the upper half the main OpenAPC data set for OA articles can be seen. The table with the additional costs that we created specifically for this test is shown in the lower half. The DOI serves as linking key there. All DOIs listed in the special

OpenAPC main data set, containing only APC costs („euro“ column)

institution	period	euro	doi	is_hybrid	publisher	journal_full_title
desy	2016	1728.73	10.1364/oe.24.021059	FALSE	Optical Society of America (OSA)	Optics Express
desy	2016	735.56	10.3390/app6090238	FALSE	MDPI AG	Applied Sciences
desy	2016	5000	10.1002/adfm.201603191	TRUE	Wiley-Blackwell	Advanced Functional Materials
desy	2016	1712.11	10.1364/oe.24.021752	FALSE	Optical Society of America (OSA)	Optics Express
desy	2016	1069.39	10.4236/jmp.2016.79088	FALSE	Scientific Research Publishing, Inc.	Journal of Modern Physics
desy	2016	1755.54	10.1364/oe.24.025169	FALSE	Optical Society of America (OSA)	Optics Express
desy	2016	1771.04	10.1364/oe.24.025582	FALSE	Optical Society of America (OSA)	Optics Express
desy	2016	912.37	10.1364/optica.3.001209	FALSE	Optical Society of America (OSA)	Optica

New, supporting openCost data set containing additional cost data types

doi	cover	colour_charges	page_charges	other	permission	publication_charges	reprint	submission_fee
10.1364/oe.24.021059	NA	NA	NA	NA	NA	NA	NA	NA
10.3390/app6090238	NA	NA	NA	NA	NA	NA	NA	NA
10.1002/adfm.201603191	NA	1495	NA	NA	NA	NA	NA	NA
10.1364/oe.24.021752	NA	NA	NA	NA	NA	NA	NA	NA
10.4236/jmp.2016.79088	NA	NA	NA	NA	NA	NA	NA	NA
10.1364/oe.24.025169	NA	NA	NA	NA	NA	NA	NA	NA
10.1364/oe.24.025582	NA	NA	1483.36	NA	NA	NA	NA	NA
10.1364/optica.3.001209	NA	NA	NA	NA	NA	NA	NA	NA

Table 3.5: DESY test data with addtitional costs table

table also occur in the OpenAPC main data set. This way, all collected additional costs can be assigned to specific entries in the main data set.

After processing the data in this way, it was analyzed in terms of the additional costs.

3.5.2 Data analysis

For the analysis of the data, these were first structured and listed according to occurrence. Simultaneously, the respective sums and medians were calculated. Table 3.6 sums up all cost types which were detected in the data.

Each article had an APC assigned so those 1933 occurrences of APCs are equivalent to the total number of articles. With a total amount of over 3.5 Mio €, APCs represent the largest share of costs by far, while the occurrence of additional costs is relatively small in comparison. With just five occurrences, the costs for preprints are the least represented. The additional costs occurring most often belong to the Others category, with 100 entries, but this is still just a small number in relation to the total number of articles. Nevertheless, even this first overview reveals an interesting aspect: The data shows that the median for additional costs can be quite high. For example, looking at the color charges, the median is 1960€, which is even higher than the APC median.

Next, we plotted this data. In doing so, we looked at the costs from three perspectives: Costs per publisher, per OA type and per institution.

Cost type	Occurrence	Median	Sum
APC	1 933	1 570	3 501 770
Cover Charge	12	1 206	14 887
Colour Charge	41	1 960	77 114
Page Charge	60	838	62 547
Other	100	27	13 230
Permission	6	125	1 098
Publication Charge	15	1 143	18 958
Reprint	5	399	2 111
Submission Fee	10	68	768

Table 3.6: Additional cost entries in the pilot data provided by several JOIN² partners.

For the analysis by publisher, we chose the ones that occurred most frequently within the 1933 articles. In this case, the publishers with more than 40 entries were picked. Figure 3.2 provides an overview of the publishers, including the actual share of the additional cost types. The light green bar, which is the largest share, is the classic APC, which has always been recorded in OpenAPC. All others are the additional costs. It can be seen that these additional costs usually do not make up a big part of the total cost of a publication in this sample but there are certain exceptions. Looking at Oxford University Press, it is apparent that the share of additional costs is quite large in comparison to other publishers, as there is a large share of color charges in relation to the total cost of publication. In addition, one can see that for OA-only publishers, like MDPI and Frontiers, additional costs tend not to be charged. So according to this sample data it seems to be most likely a phenomenon associated with Hybrid OA publishing.

This hypothesis is confirmed by the next analysis(see Figure 3.3). Here we analyzed the data by OA Type, i.e., whether the articles were published in a Gold OA or Hybrid OA journal. The visualization by OA types illustrates that in fact those additional cost types mostly occur with Hybrid OA.

Finally, an analysis by institution was conducted, which is shown in Figure 3.4. However, it must be noted that the results do not provide much added value in terms of additional costs. While it can be seen that most of the additional cost types occur at the FZJ, this needs to be put into perspective since this institution has the largest share of costs anyway, compared to DESY and GSI.

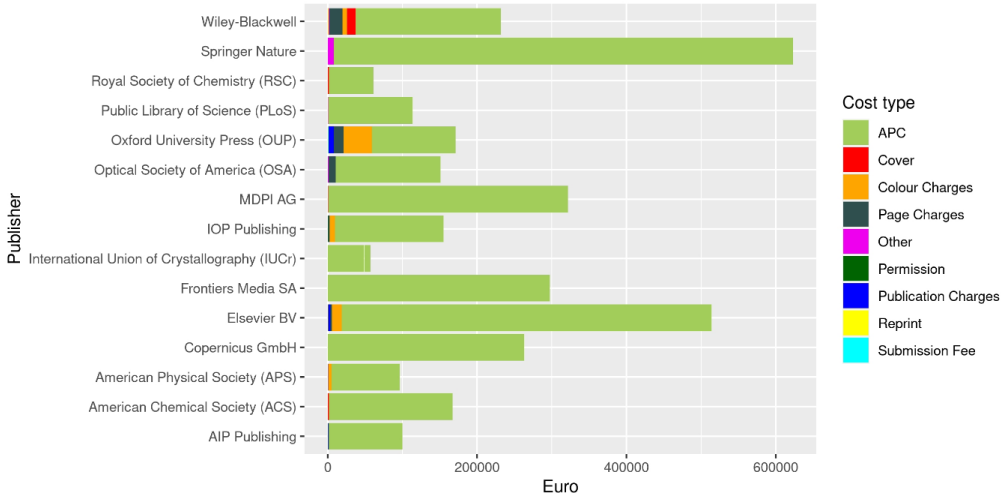

Figure 3.2: Analysis: Costs per publisher

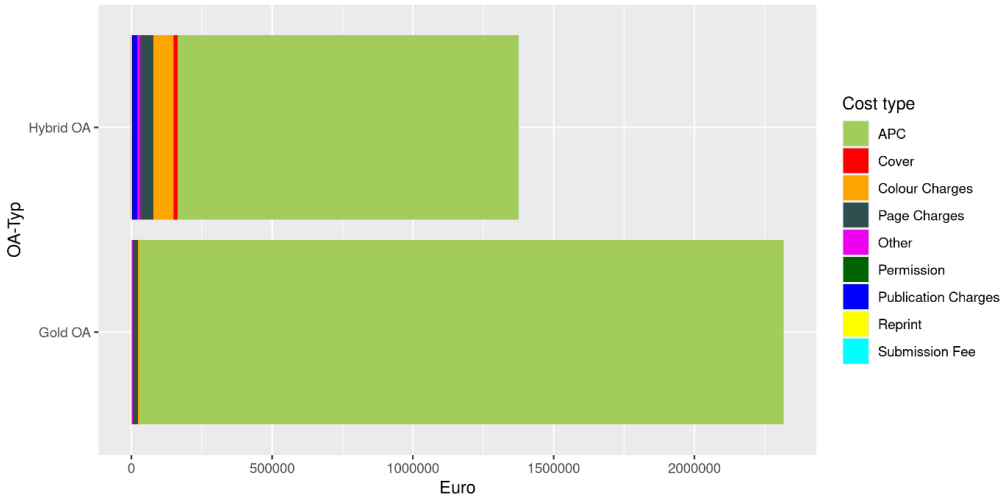

Figure 3.3: Analysis per OA type

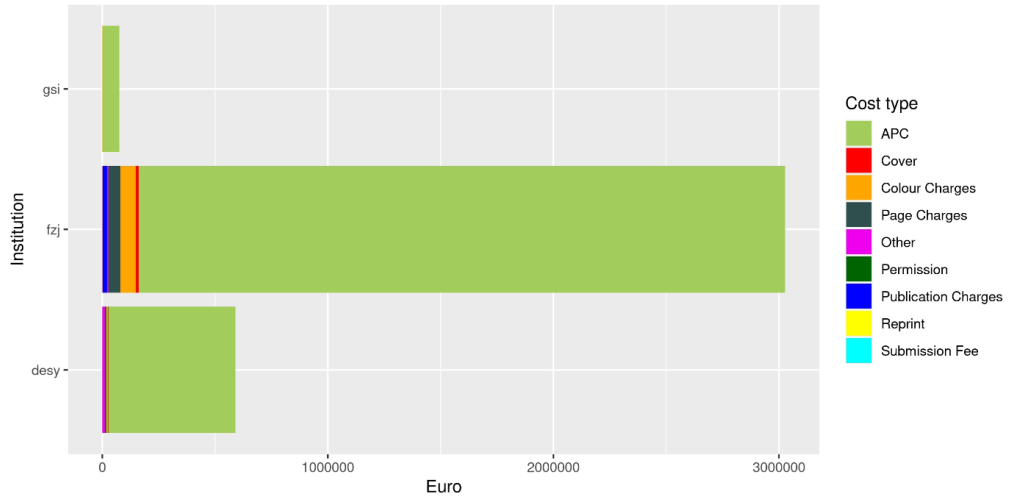

Figure 3.4: Analysis per institution

3.6 Conclusion

The analysis conducted in the preliminary pilot test has shown that the aggregation of additional cost types within the OpenAPC framework is possible without any serious problems. The proposed approach of storing these extra costs in an additional table and linking this data to the main dataset via DOI does not interfere with the established OpenAPC data, meaning that long-term consistency is not affected.

The data indicates that additional cost types occur more often with certain publishers and also more often in the field of hybrid OA publishing. In this context it should be noted, however, that the data is probably somewhat biased. Less than 2,000 articles were available for this pilot project. The three institutions from which the data was harvested are research institutes, so universities – which usually have a big share in the area of academic publishing – are not represented here. In addition, there is also possibly a subject bias in the data, as all institutions are STEM-focused. Thus, publications from the humanities and social sciences are less likely to be present in the data.

However, despite the above mentioned bias, this test run provides some interesting preliminary evidence that extending the OpenAPC infrastructure by adding this additional cost data may complete the picture on fee-based OA publishing and thus contribute to improved cost transparency.

References

[19] JOIN². *JOIN² - Just anOther INvenio INstance To Join*. en. URL: https://join2.de/ (visited on 2022-12-08) (cit. on pp. 11, 12, 38, 103).

[22] JOIN². *JOIN² Instanzen*. 2022. URL: https://join2.de/Main/Instanzen (visited on 2022-12-06) (cit. on pp. 12, 22, 38).

[43] Gernot Deinzer et al. "Including Open Access Information in institutional repositories". In: CERN Workshop on Innovations in Scholary Communication, Geneva (Switzerland), 21 Jun 2017 - 23 Jun 2017. June 21, 2017. DOI: 10.3204/PUBDB-2017-05961 (cit. on pp. 16, 37).

[45] openCost. *openCost*. URL: https://www.opencost.de (visited on 2022-12-13) (cit. on pp. 16, 37, 46).

[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, 112).

[54] The Open Knowledge Foundation. *Open Database License (ODbL) v1.0*. Nov. 30, 2022. URL: https://opendatacommons.org/licenses/odbl/1-0/ (visited on 2022-11-30) (cit. on p. 32).

[55] *INTACT project*. 2022. URL: https://gepris.dfg.de/gepris/projekt/274983645?language=en (visited on 2022-11-30) (cit. on p. 32).

[56] OpenAPC. *APC data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/apc_de.csv (visited on 2022-11-30) (cit. on p. 32).

[57] OpenAPC. *TA data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/transformative_agreements/transformative_agreements.csv (visited on 2022-11-30) (cit. on p. 32).

[58] OpenAPC. *BPC data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/bpc.csv (visited on 2022-11-30) (cit. on p. 32).

[59] OpenAPC. *OpenAPC GitHub repository*. 2022. URL: https://github.com/OpenAPC/openapc-de (visited on 2022-11-30) (cit. on p. 33).

[60] OpenAPC. *OpenAPC first release*. 2022. URL: https://github.com/OpenAPC/openapc-de/tree/v1.0.0 (visited on 2022-11-30) (cit. on p. 33).

[61] OpenAPC. *OpenAPC README file*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/README.md (visited on 2022-11-30) (cit. on p. 33).

[62] OpenAPC. *OpenAPC data schemas*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/schema (visited on 2022-12-06) (cit. on p. 34).

[63] Dirk Pieper and Christoph Broschinski. "OpenAPC: a contribution to a transparent and reproducible monitoring of fee-based open access publishing across institutions and nations". In: *Insights the UKSG journal* 31 (2018), p. 39. DOI: 10.1629/uksg.439 (cit. on pp. 34, 47).

[64] OpenAPC. *OpenAPC cost definition*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/Data-Submission-Handout#definition-of-costs (visited on 2022-11-30) (cit. on p. 35).

[65] OpenAPC. *OpenAPC handout on DEAL data*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/Handreichung-Dateneingabe-Transformationsvertr%C3%A4ge-%28DEAL-Wiley-und--Springer-Nature%29-ab-Berichtsjahr-2020 (visited on 2022-11-30) (cit. on p. 35).

[66] OpenAPC. *OpenAPC blog*. 2022. URL: https://openapc.github.io/ (visited on 2022-11-30) (cit. on p. 35).

[67] OpenAPC. *OpenAPC treemaps*. 2022. URL: https://treemaps.openapc.net/ (visited on 2022-11-30) (cit. on p. 36).

[68] OpenAPC. *OpenAPC OLAP server*. 2022. URL: https://olap.openapc.net/ (visited on 2022-11-30) (cit. on p. 36).

[69] Bianca Schweighofer. "Projekt openCost auf dem 8. Bibliothekskongress 2022". de. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* (2022). Artwork Size: 1-10 Seiten Publisher: o-bib. Das offene Bibliotheksjournal / Herausgeber VDB, pp. 1–10. DOI: 10.5282/O-BIB/5862 (cit. on pp. 37, 96).

4

Information Budget: 12 + 6 + 8 = 10

Bernhard Mittermaier[1]
Forschungszentrum Jülich, Central Library

DOI: 10.3204/DESY-PROC-2023-01/04

Abstract

The transformation of scientific publishing into Open Access is the impetus for a holistic view of the expenditure both for access to scientific literature and for publishing. What happens in a Publish & Read contract on a small scale must also be the approach overall, in accordance with the demands from funders (German Research Foundation, DFG) and policy advice (German Science and Humanities Council, WR). This article presents an information budget in its entire breadth as well as in the details necessary for practical implementation, both for the expenditure side and for the sources of funding. Included is a workshop report which demonstrates the practical suitability of the concept using the example of Forschungszentrum Jülich.

Keywords: information budget; open access; transformation; total cost of publication; publication costs; acquisition budget

[1] https://orcid.org/0000-0002-3412-6168

4.1 Introduction[2]

In Germany, the topic of information budgets has been a hot topic in academic institutions and especially academic libraries, at the latest since the German Council of Science and Humanities (Wissenschaftsrat) called for the establishment of information budgets on p. 9 in its Recommendations on the Transformation of Academic Publishing: Towards Open Access [70]

> The Council therefore recommends that scientific institutions record all components of their information budget and balance possible savings on the acquisition side as well as possible sources of income against rising expenditure on publication services.

The topic was subsequently addressed in a series of workshops organized by BMBF's open-access.network project[3]. The DFG project openCost [45] addresses aspects of the topic at the operational level and held a hands-on lab and expert session at the German Library Congress in Leipzig [71] and an expert workshop in October 2022 [72]. In addition to this, the DFG project Transform2Open [49] deals with the conceptual aspects of transformation.

This article attempts to describe all components of an information budget and to mirror this conceptual approach with a real example.

4.2 The concept of an information budget

4.2.1 Emergence

To finance Gold Open Access publication fees, the first academic institutions set up publication funds in the first decade of the 21st century [73]. At the beginning of the second decade, Ralf Schimmer called for the "re-contextualisation of acquisition budgets", i.e. the integral consideration of publication and subscription expenditure and the transformation of acquisition expenditure into publication expenditure [74]. In an international discussion, this integral view was often

[2] This article is based on a presentation given at the expert workshop openCost: The Road to Publication Cost Transparency, Hamburg, 05.–07.10.2022 https://juser.fz-juelich.de/record/915882 and on a publication in German language [10].

[3] Workshops Budget Development in the Context of Open Access Transformation, <https://open-access.network/fortbilden/thematische-workshops/workshops-budgetentwicklung-im-kontext-der-open-access-transformation>, accessed: 13.12.2022.

called for, often under the catchphrases *total cost of ownership* [75, 76] or *total cost of publication* [77, 78]. In 2015, public accounting of open access (OA) publication fees began as part of the OpenAPC project [63] and in 2017, the public presentation of all publication fees and acquisition costs by Forschungszentrum Jülich in its Open Access Barometer [79]. On a much larger scale, university institutions in the UK published the subscription costs paid by them [80]. The Open Access Working Group of the Digital Information priority initiative of the Alliance of Science Organizations in Germany recommended such an approach to all scientific institutions [81]. Furthermore, the Ad Hoc Working Group on Open Access Gold of the Priority Initiative recommended in 2016 [82]

> that libraries create the necessary structural conditions for this and jointly focus on publication funds and acquisition budgets as an integral part of their budget planning and management. Institutions with multi-tiered library systems, in the context of which subscriptions are jointly funded by several parties, should seek to centralize funds to ensure monitoring and control of financial flows between publishers and academic institutions.

For a long time, however, these appeals found little resonance, all the more so with regard to a public presentation of the results, which even led to subversive approaches [83].

Now, at the beginning of the 2020s, the topic is being discussed more broadly as information budget. An early treatment was carried out by Heinz Pampel in 2019 [12]; in 2022 he deepened his considerations [84]. On a very practical level, discussions arose (usually without actually mentioning the term information budget) when institutions participating in DEAL contracts were asked for "voluntary additional payments" as part of the True-Up in autumn 2020 [85] and many institutions did not comply with this request. While there was a *de facto* saving of Hybrid OA costs, these were funded directly by academia prior to the DEAL contracts, which meant that the funds were not available in the library budgets for post-payments. Great momentum was created at the beginning of 2021 by the German Research Foundation's (DFG) funding programme Open Access Publication Funding. Its overarching aim is "to create structural adjustment in financial flows, thus enabling the transformation of open access and improving transparency regarding the costs for open access publication of research results." [14] Applicant institutions must outline how they intend to achieve this aim, which ultimately entails the creation of information budgets. Then, at the beginning of 2022, the

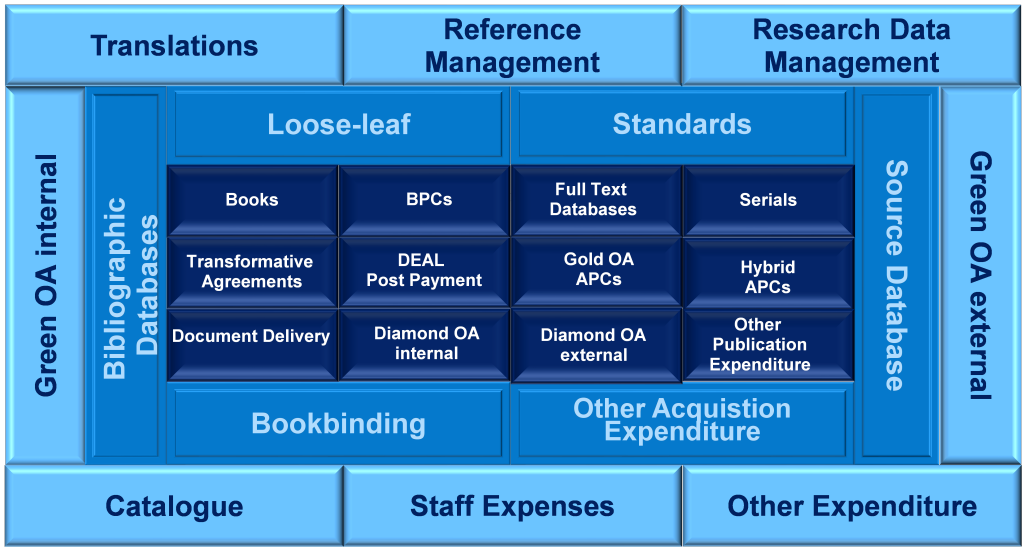

Figure 4.1: Shell model of the information budget

aforementioned Recommendations on the Transformation of Academic Publishing: Towards Open Access [70] were published by the German Council of Science and Humanities. These recommendations also take a holistic approach: It is not *only* about interdependencies between journal subscriptions and publication fees for journal articles in OA, but fundamentally about all expenditure (and income) for both the procurement of information and for publishing.

In the following, an information budget is conceptualized on the expenditure side (subsection 4.2.2) and the income side (subsection 4.2.3).

4.2.2 The expenditure side

The proposal for an information budget in Figure 4.1 is very comprehensive with a total of 26 parameters. It is modelled in the form of different shells or layers for two reasons. Firstly, the large number of parameters should not deter people from tackling the task. An institution does not have an information budget only when the last parameter listed here has been recorded. Rather, an information budget can also begin with partial aspects, particularly with the parameters at the core. Secondly, it is structured based on the (presumed) possibilities and requirements from the perspective of German institutions.

- The core (dark blue) contains parameters required for the DFG funding programme Open Access Publication Funding [14] as well as other cost categories related to books and journals. This core is of utmost importance for the transformation process.
- The first shell (blue) contains all other acquisition costs.
- The second shell (light blue) includes other cost categories including staff expenses, as well as the costs for the library catalogue and research data management.

In the process, not only sum values should be recorded, but also each cost category for each publisher. If this is not already the case or can easily be facilitated, the costs have to be allocated to the publishers in the same proportion as the costs for the differentiated cost categories (cf. section 4.3). Although neither the German Council of Science and Humanities nor the Priority Initiative explicitly call for publisher-specific data collection, it is essential to achieve the goal of cost reduction. The DFG programme Open Access Publication Funding [14] calls for this differentiation by publisher "where possible".

Not only the allocation to the three areas, but also the partial or complete (non-) inclusion of a cost category in the information budget is undoubtedly debatable. The deciding factors were the questions:

1. **Does the cost category enable the production or dissemination of a publication?**

   There are grey areas here, e. g. translations. If a scientific publication is translated or proofread by an internal or external language service before submission, then these costs are taken into account. However, if this work is carried out by the authors or their colleagues, the costs are not taken into account. According to the wording, the question would also cover the costs of the research itself, but this is of course not purposeful and therefore not done.

2. **Does the cost category enable the purchase or provision of media?**

   The costs of providing the literature, e. g. the costs of the reading room, could also be included. However, this was dispensed with.

3. **Does the cost category support open access?**

This criterion was included in order to account for expenditure for Diamond open access that does not directly benefit publications by your own institution.

Cost catagories are then included in Figure 4.1 when at least one of these questions can be answered in the affirmative.

4.2.2.1 Notes on individual facets

Core: Books and journals

- Books, Journals, Standards and Loose-leaf include both printed and electronic editions, individual purchases, as well as packages, temporary licences, and permanent purchases.
- Full-text databases are aggregated databases that contain full texts of books and/or journals and thus replace them. Other databases are included in the first shell.
- Transformative agreements combine aspects of reading and writing and therefore form a category of their own. This can also include Subscribe2Open.
- Publication fees are differentiated into DEAL post-payments, Gold OA APCs, Hybrid OA APCs, and Other publication expenditure. The latter includes, for example, colour charges, cover charges, page charges, submission fees, printing cost subsidies, and permissions (e. g. Copyright Clearance Centre). DEAL post-payments are separate from transformative agreements because they only contain a publication component.
- Document delivery includes the costs of receiving interlibrary loan, document delivery, and pay-per-view.
- Diamond OA internal is the expenditure for the university's own publishing house etc.; Diamond OA external is the cost of co-funding Diamond OA that is practised at other institutions.

First shell: Other media

- Bibliographic databases mainly contain references to other (full-text) sources, e. g. Dimensions, Scopus, Web of Science. Source databases, such as

SciFinder mainly contain independent content, but no journals or books (these are in full-text databases at the core).

- Bookbinding includes both the expenditure for externally contracted bookbinding work and the expenditure for any in-house bookbinding.
- Other acquisition expenditure includes, for example, AV media, newspapers, and media treated as consumables.

Second shell: Other expenditure

- Green OA internal includes expenditure that make Green OA in one's own institution possible, e. g. costs of an institutional repository. Green OA external includes payments for Green OA at other institutions, e. g. for arXiv.
- Reference management is the cost of commercial reference management software.
- Staff expenses refer to the personnel costs for all persons directly involved in the aforementioned tasks, if necessary on a *pro rata* basis if they also carry out other activities. Library staff with overlapping tasks (e. g. management, administration, …) are taken into account by calculating the share of the first-mentioned persons in relation to all library staff.

4.2.3 The revenue side

Sources of funding include both the library and the faculties/divisions/institutes/-working groups etc., hereinafter referred to as institutes [86]. A distinction can be made between own resources and third-party funding, but also between the acquisition budget and the publication fund, although the extent to which the institutes actually plan these budget areas is unclear. Other budget items also play a role in the library's financial resources, but these are no longer differentiated and are all subsumed under Other budget sources. Thus, in a somewhat simplistic representation, a total of ten facets emerge (Table 4.1).

Of course, these sources only count as part of an information budget insofar as they serve to finance acquisition costs and publication fees and not, for example, the entire financial resources of the institutes.

Although it is in the library's own interest that funding sources from the institutes are maximized to save library funds, for the institutes, the opposite is

		Own resources	**Third-party funding**
Institutes	Acquisitions	Used in cases where third-party funding is unavailable and the library does not cover the costs. Generally, this is institute libraries' main source of funding.	Used when available and in cases where the library does not cover the costs. Often relates to material costs for literature in third-party projects.
	Publication fund	Used in cases where third-party funding is unavailable and the library does not cover the costs. Often involves Hybrid OA fees in the wild and publication charges for closed access publications.	Used when available and in cases where the library does not cover the costs. Often involves Hybrid OA fees in the wild and publication charges for closed access publications.
Libraries	Acquisitions	This is how the bulk of acquisition expenditure is financed.	Examples include specialist information services operated by the library and prepaid consortium licences for an entire consortium.
	Publication fund	The general rule for publication fees paid by the library.	Funding within the framework of the DFG Open Access Publication Funding programme is an example.
	Other budget sources	The general rule for financing expenditure from the second shell of the information budget model.	Third-party funding is used to finance expenditure from the second shell of the information budget model in particular, e. g. financing a library catalogue from consortia funds.

Table 4.1: Facets of funding sources.

true. The idea of the information budget is to overcome these perceptions and look at the overall picture of relevant expenditure and its sources of funding. In addition to the naturally obvious goal of both library and institutes acquiring more third-party funding, the goal must be to fully utilize existing third-party funding from institutes and not let it lapse. However, distribution struggles can exist not only between library and institutes, but also within the library if publication funding and acquisition budgets are two strictly separate entities.

4.3 Forschungszentrum Jülich's workshop report

The Central Library of Forschungszentrum Jülich began detailed cost monitoring at an early stage, the nucleus of which was an ERM system developed in-house [87]. Since 2016, all forms of expenditure for journals have been recorded and published in the OA Barometer [88]. An information budget has thus been realized for this sub-sector. In Table 4.2, the parameters from Figure 4.1 are shown as a proportion of the expenditure for Jülich as a whole in 2021. This also shows where expenditure already occurs at publisher level, where this is the aim, and where a permanent distribution to publishers is made according to the shares in the other positions.

	Share 2021	**Comment**	**Publisher level**
Core: Books and journals			
Books	5.14 %		Desideratum
BPCs	0.00 %	Not applicable	✓
Full-text databases	0.00 %	Not applicable	Pro rata
Journals	8.74 %		✓
Transitive agreements	10.59 %		✓
DEAL post-payments	3.31 %	Payment for 2021 in 2022	✓
Gold OA APCs	12.77 %		✓
Hybrid APCs	6.26 %	Decentralized costs[4]	✓
Document Delivery	1.60 %		Desideratum
Diamond OA internal	0.05 %		✓
Diamond OA external	0.05 %		✓
Other publication expenditure	2.16 %	Decentralized costs	
First shell: Other media			
Bibliographic databases	6.52 %		✓
Source database	5.93 %		✓
Loose-leaf	0.16 %		Desideratum
Standards	0.59 %		Desideratum
Bookbinding	0.04 %		Pro rata
Other acquisition expenditure	0.61 %		Desideratum
Second shell: Other expenditure			
Green OA internal	0.00 %		Pro rata
Green OA external	0.41 %		Pro rata

[4] Expenditure in the cost categories Hybrid APCs and Other publication expenditure is processed by the Central Library, but is charged to the client's cost centre. Coverage is not complete, but is estimated at over 95%.

	Share 2021	Comment	Publisher level
Translations	0.78 %		Pro rata
Reference management	0.35 %		✓
Research data management	0.08 %		Pro rata
Library catalogue	0.40 %		✓
Staff expenses	33.18 %	47% of the library's staff expenses were included	Pro rata
Other expenditure	0.30 %	Plagiarism software; scanner maintenance; SFX; EZB fee	Pro rata

Table 4.2: Distribution of Forschungszentrum Jülich's expenditure across the cost categories of the information budget. The column Publisher level refers to issues that can be assigned to a specific publisher/provider. The ticks ✓ indicate that the relevant information is available. Desideratum means that the expenditure is known as a sum total but is not currently resolved at publisher level. This will be implemented in 2023. Pro rata means that the total costs are not permanently levied at publisher level, but are allocated to the publishers on a pro rata basis in the same proportion as the other cost categories are distributed among the publishers.

Figure 4.2 is a graphical representation of the expenditure of Forschungszentrum Jülich within the framework of the information budget. The largest single item is staff expenses, followed by Gold OA APCs. These are the dominant types of expenditure in the core area of the information budget, followed by transformative agreements. Classic subscription fees only follow in fourth place at Forschungszentrum Jülich. Hybrid OA APCs in the wild, which have to be paid by the authors' institutes, account for a remarkably high proportion. From this, it can be concluded that there is a high intrinsic motivation for open access publications.

The expenditure categories that were not recognizable in Figure 4.2 (share < 1 %) are shown enlarged in Figure 4.3 It should be noted that the expenditure for Diamond OA is very low internally, because this analysis only accounts for the expenditure for external hosting and does not account for staff expenses. Not shown are Green OA internal, BPCs and full-text databases, all of which account for 0 %. For Green OA internal, this is the effect of the separately presented staff expenses. BPCs and expenses for full-text databases were indeed not incurred.

A good 84 % of Forschungszentrum Jülich's information budget in 2021 (Figure 4.4) comes from the library's own resources. A total of 15.4 % was funded by the institutes; the largest share of this was publication fees financed through

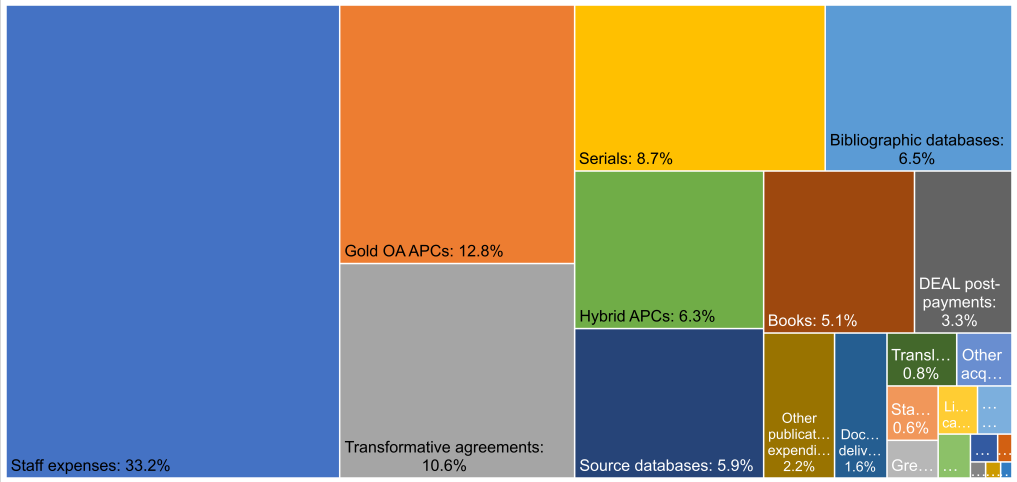

Figure 4.2: Expenditure of Forschungszentrum Jülich within the framework of the information budget.

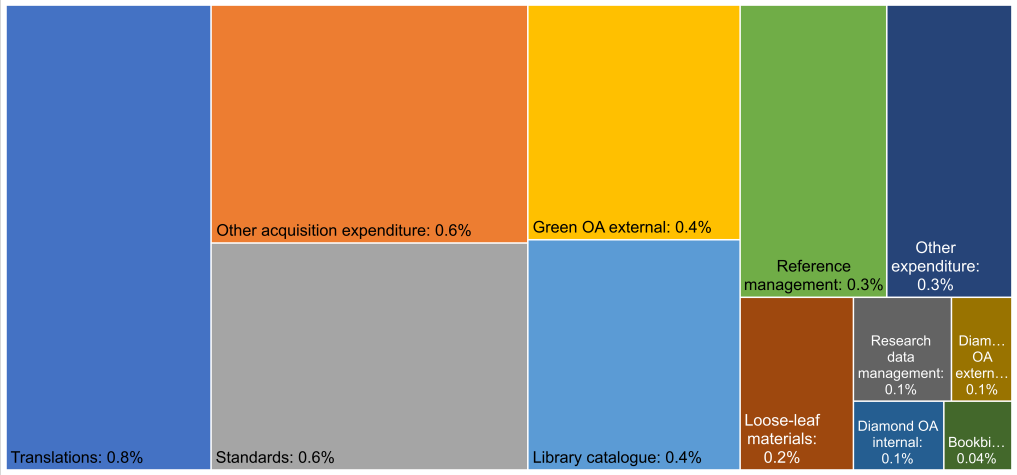

Figure 4.3: Expenditure of Forschungszentrum Jülich within the information budget with a share of less than 1 %.

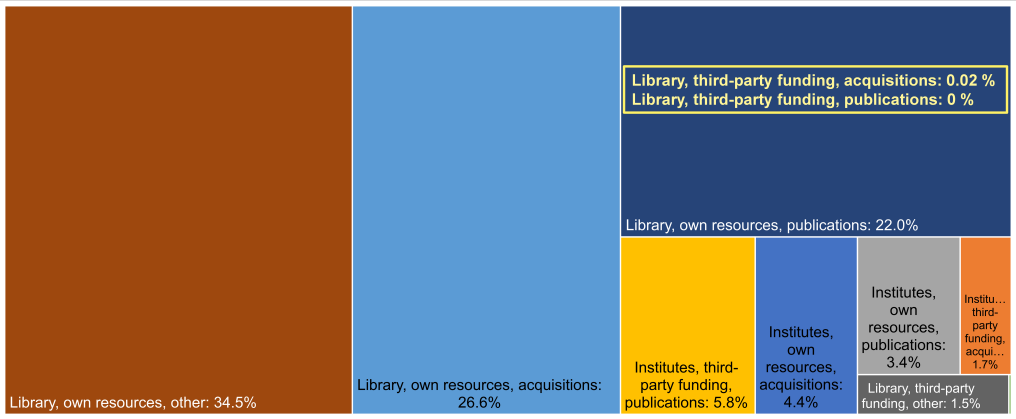

Figure 4.4: Sources of funding for Forschungszentrum Jülich's information budget in 2021.

third-party funding (5.8 %). The largest part of the institutes' expenditure (13.4 %) was handled by the Central Library; the library only became aware of 2.0 % of the information budget expenditure (corresponding to 3.2 % of the acquisition costs and publication fees in the narrower sense) in the course of searches in the SAP accounting software.

At the Central Library, third-party funding only plays a minor role in the information budget, in the form of overhead funding within a DFG-funded transformative agreement (acquisition) and an NFDI4Ing-funded position in research data management. Third-party funding will only become relevant in publication expenditure from 2022 onwards (DFG programme Open Access Publication Funding).

The ten segments of Figure 4.4 are grouped in Figure 4.5 in such a way that the proportions of those responsible for financing, the sources, and the budgets can be seen.

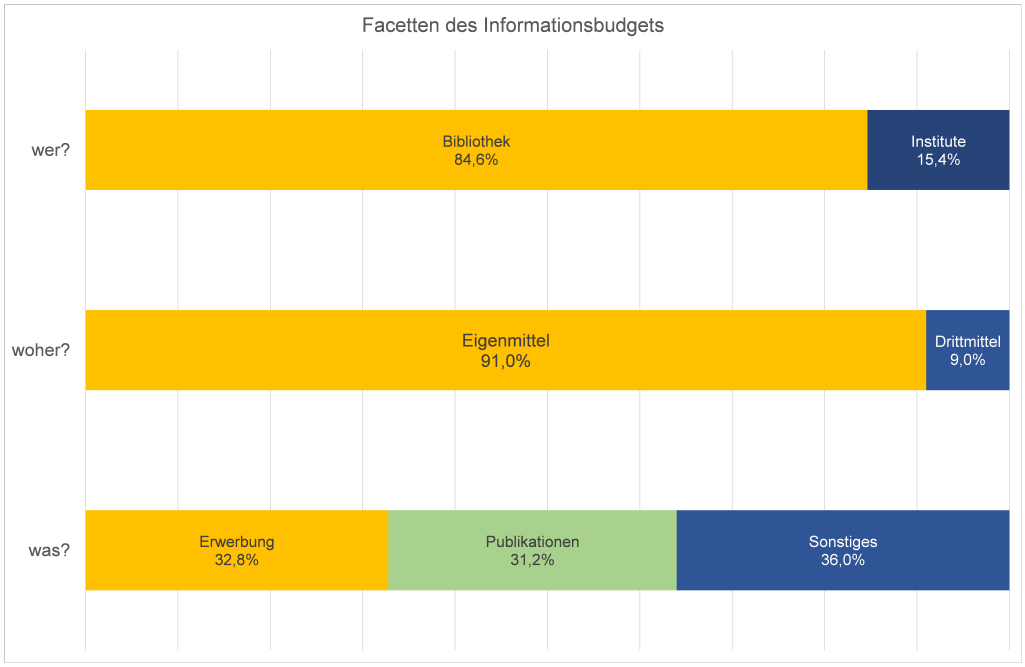

Figure 4.5: Facets of the Forschungszentrum Jülich's information budget for 2021. Transformative agreements were divided into acquisition and publication according to the publisher's details (e. g. by VAT rate).

4.4 Discussion

The preparation of an information budget (or more precisely: the aggregation of expenditures and their funding sources) is a challenging task. While all the figures are generally available in an *ex post* view, all of the necessary information related to the figures is often only available indirectly. Did the institute finance this book from its own resources (basic funds) or using third-party funding? Which publisher does this payment to the Copyright Clearance Center concern? What literature was acquired outside of the official procurement channels? These and similar questions are always asked when compiling the data corpus. Another problem is that staff expenses represent the largest item but they are probably the most imprecise to determine. Allocating them to the information budget is only easy if a person undoubtedly works entirely in this area. As soon as proportions have to be estimated, the calculation inevitably becomes imprecise. In addition, the approach of allocating categories of expenditure on a *pro rata* basis where this cannot be avoided has reached its limits in the case of staff expenses. This is particularly clear in the case of Green OA internal, where no costs were incurred[5], which is why no staff expenses are charged. However, the real staff expenses for the repository comprise about three full-time positions.

Given 26 categories – distributed over three shells of 12, 6, and 8 categories – the expenditure side of the information budget is already very detailed, even if additional steps of detail are conceivable. However, it may also be too detailed for practical use; at Forschungszentrum Jülich, for example, 15 of the 26 categories account for less than 1 % of total expenditure. Should this very skewed distribution also occur in other institutions, a readjustment of the model would have to be considered, e. g. by combining some categories. In any case, it will be interesting to be able to look at the development over time after this point measurement concerning the year 2021 by repeating it in subsequent years.

On the revenue side, 12 categories result very stringently from, firstly, the differentiation between library and faculties/institutes, secondly, the differentiation between own resources and third-party funding, and thirdly, the differentiation between acquisition budget, publication budget, and other budget. However, faculties/institutes normally only have their own staff budget in two-tier library systems, which is the main part of Other. For Forschungszentrum Jülich, the categories Institute/Other/Own resources and Institute/Other/Third-party funding

[5]This is also not quite correct in itself because energy costs were not taken into account and server costs were not depreciated over their useful life.

are missing. This results in 10 categories and hence the equation in the title of this paper.

Finally, it should be emphasized that the concept was developed from a German perspective, cf. the category DEAL post-payments for example, which is only relevant in Germany. However, after minor adjustments, the ability to apply it to an international context is possible and desired.

4.5 Acknowledgements

I would like to thank my colleagues Kai Pelzer and Irene Barbers for their support in obtaining the data, Hilde Dobbelstein for producing Figure 4.1 the document, and the Language Services of Forschungszentrum Jülich for linguistic support. This work was funded by the German Research Foundation (DFG) — 505575192.

References

[10] Bernhard Mittermaier. "Informationsbudget". In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 9.4 (2022), pp. 1–17. DOI: 10.5282/o-bib/5864 (cit. on pp. 4, 46).

[12] Heinz Pampel. *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten*. de. Arbeitspapier. 2019. DOI: 10.2312/OS.HELMHOLTZ.006 (cit. on pp. 6, 17, 47, 72, 75).

[14] Deutsche Forschungsgemeinschaft. *Funding Programme: "Open Access Publication Funding"*. URL: https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/open_access_publikationskosten/ (visited on 2022-12-13) (cit. on pp. 8, 16, 21, 22, 47, 49).

[45] openCost. *openCost*. URL: https://www.opencost.de (visited on 2022-12-13) (cit. on pp. 16, 37, 46).

[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 (cit. on pp. 21, 46, 95, 98, 121).

[63] Dirk Pieper and Christoph Broschinski. “OpenAPC: a contribution to a transparent and reproducible monitoring of fee-based open access publishing across institutions and nations”. In: *Insights the UKSG journal* 31 (2018), p. 39. DOI: 10.1629/uksg.439 (cit. on pp. 34, 47).

[70] German Science and Humanities Council. *Recommendations on the Transformation of Academic Publishing. Towards Open Access*. Position Paper Drs. 9477-22_en. Jan. 2022. DOI: 10.57674/0gtq-b603 (cit. on pp. 46, 48, 95).

[71] openCost. *openCost auf dem Bibliothekskongress: Ergebnisse des Workshops*. URL: https://www.opencost.de/allgemein/ergebnisse-des-workshops/ (visited on 2022-12-13) (cit. on p. 46).

[72] *Workshop “openCost: the road to publication cost transparency”*. URL: https://www.opencost.de/allgemein/opencost-workshop/ (visited on 2022-12-13) (cit. on p. 46).

[73] Anita Eppelin et al. “Umgang mit Open-Access-Publikationsgebühren – die Situation in Deutschland in 2010”. In: *GMS Medizin - Bibliothek - Information* 12.1-2 (2012). DOI: 10.3205/MBI000240 (cit. on p. 46).

[74] Ralf Schimmer. “Open Access und die Re-Kontextualisierung des Bibliothekserwerbungsetats”. In: *Bibliothek Forschung und Praxis* 36.3 (2012), pp. 293–299. DOI: 10.1515/bfp-2012-0038 (cit. on p. 46).

[75] Stuart Lawson. “’Total cost of ownership’ of scholarly communication. managing subscription and APC payments together”. In: *Learned Publishing* 28.1 (2015), pp. 9–13. DOI: 10.1087/20150103 (cit. on p. 47).

[76] Irene Barbers, Nadja Kalinna, and Bernhard Mittermaier. “Data-Driven Transition: Joint Reporting of Subscription Expenditure and Publication Costs”. In: *Publications* 6.2 (2018). DOI: 10.3390/publications6020019 (cit. on p. 47).

[77] Andrew Gray. “Considering Non-Open Access Publication Charges in the “Total Cost of Publication””. In: *Publications* 3.4 (2015), pp. 248–262. DOI: 10.3390/publications3040248 (cit. on p. 47).

[78] Stephen Pinfield, Jennifer Salter, and Peter A. Bath. "The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions". In: *Journal of the Association for Information Science and Technology* 67.7 (2016), pp. 1751–1766. DOI: 10.1002/asi.23446 (cit. on pp. 47, 72, 75, 87).

[79] Bernhard Mittermaier and Irene Barbers. "Zeitschriftenmonitoring". In: *Bibliotheken. Wegweiser in die Zukunft. : Projekte und Beispiele*. Ed. by Erda Lapp et al. Berlin: Simon Verlag für Bibliothekswissen, 2021, pp. 346–359. DOI: http://hdl.handle.net/2128/22746 (cit. on pp. 47, 75).

[80] Stuart Lawson and Ben Meghreblian. "Journal subscription expenditure of UK higher education institutions". In: *F1000Research* 3 (2015), p. 274. DOI: 10.12688/f1000research.5706.3 (cit. on p. 47).

[81] Arbeitsgruppe Open Access in der Schwerpunktinitiative „Digitale Information" der Allianz der Deutschen Wissenschaftsorganisationen. *Open Access gestalten, Transparenz schaffen! Appell zur Offenlegung von Publikationsgebühren*. 2016. DOI: 10.2312/ALLIANZOA.013 (cit. on p. 47).

[82] Christoph Bruch et al. *Empfehlungen zur Open-Access-Transformation. Strategische und praktische Verankerung von Open Access in der Informationsversorgung wissenschaftlicher Einrichtungen*. Recommendations. Ad-hoc-AG Open-Access-Gold im Rahmen der Schwerpunktinitiative „Digitale Information" der Allianz der deutschen Wissenschaftsorganisationen, 2016. DOI: 10.3249/ALLIANZOA.011 (cit. on p. 47).

[83] Christian Gutknecht. "Transparenz von Subskriptionskosten in der Schweiz". In: *027.7 Zeitschrift für Bibliothekskultur* 4.1 (2016), pp. 26–32. DOI: 10.12685/027.7-4-1-103 (cit. on p. 47).

[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 (cit. on pp. 47, 75, 94, 95, 121).

[85] DEAL Operations. *Here's the deal: Reduce and restructure costs*. URL: https://web.archive.org/web/20221128222942/https://deal-operations.de/en/here-is-the-deal/restructure-costs (visited on 2022-12-13) (cit. on p. 47).

[86] Gali Halevi and Samantha Walsh. “Faculty Attitudes Towards Article Processing Charges for Open Access Articles”. In: *Publishing Research Quarterly* 37.3 (2021), pp. 384–398. DOI: 10.1007/s12109-021-09820-x (cit. on p. 51).

[87] Florian Aumeier and Ingrid Heinen. “Inhouse-Lösung für das Jülicher Electronic Resource Management System”. In: *Bibliotheksdienst* 41.3 (2007). DOI: 10.1515/bd.2007.41.3.322 (cit. on p. 54).

[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, 120).

5

Value Assessment of License Agreements and Publication Cost from a Norwegian Perspective

Jens H. Aasheim[1]
Sikt - Norwegian Agency for Shared Services in Education and Research

DOI: 10.3204/DESY-PROC-2023-01/05

Abstract

As the national Open Access coordinator and managing the Norwegian license consortium, Sikt processes a wide variety of data associated with the cost of publishing and agreements. The following will share experiences on acquiring, handling and using cost data.

5.1 Introduction

Presented here is a brief overview of some of the past efforts and current Norwegian activities on the financial aspects of Open Access and research publishing. It is worth mentioning that at the time of writing, a larger and more comprehensive project on this topic is under way. Similar to the French study of the evolution of APC cost and electronic subscriptions [89]. However, the Norwegian project plans on extending to an institutional level as well as the national level. The project aims to establish better and more extensive cost estimates of both Open Access

[1] https://orcid.org/0000-0003-4022-8424

in Norway, in addition to the wider context of scholarly publication. Including read-only agreements, publishing agreements, APC cost, publication repositories, Diamond OA journals, etc.

Seen in an international context, Norway is a relatively centralised country: There is one main research funder, the Research Council of Norway. The Norwegian government issued national goals and guidelines for Open Access in 2017 [90]. There is a national CRIS (Current Research Information System), as a part of the governments result-based redistribution of research funding [91]. A national publication repository is currently under development, in connection with the national CRIS. And there is also a consortium negotiating and managing license agreements for most public research institutions (universities, university colleges, public research institutes, hospitals, etc.).

Sikt plays an important part in the Norwegian scholarly communications landscape. As well as hosting the national CRIS, and the upcoming joint national publication repository/CRIS. It is also both the national Open Access coordinator, manages the license agreement consortium and the national library consortium. A consequence of this central role is that Sikt also acts as a data hub for lots of relevant information. Gathering and receiving data from a plethora of different sources. Sikt is therefore in a unique position to assess, monitor and analyse both the volume and cost of Norwegian open research publication.

5.2 Transition to Open Access

In 2021 Sikt, in collaboration with Universities Norway, wrote a report on the Norwegian transition to Open Access [92].

As part of the report, Sikt sent out a survey to map how much Norwegian research institutions spend on research journals and OA publishing. In the survey institutions reported what they spend on subscriptions and agreements outside the consortium, as well as APCs and other open access related expenditures. Additionally, Sikt has direct knowledge of agreements negotiated and managed on behalf of its members through the license consortium. When combined, these elements provide a relatively complete picture of the recent historic direct costs.

Brief summary for 2019 (approximate numbers):

- The cost of the consortium agreements is almost 310M NOK[2]. Covering traditional read-only subscriptions, transformative agreements (publish&read)

[2] 100 NOK $\approx$ 9 EUR

and OA-agreements. I.e., read access and publication costs.

- 115M NOK was spent on subscriptions and agreements outside the consortium. Typically, small bundles or single journal subscriptions.
- The institutions spent another 25M NOK on publishing in fully OA journals.
- It is also estimated that an additional 30(+)M NOK was spent on Hybrid OA publishing.
- In total, Norwegian institutions spent over 480M NOK on journal subscriptions and open publishing in 2019.

Except for the consortium agreements, most of the numbers are to be regarded as approximations. Some are relatively accurate, while others are estimates and based on educated guesswork.

The most challenging part of such a cost mapping endeavour is determining the number of open publications, outside consortium agreements, to attribute to an institution and estimate the subsequent APC costs. As part of the Sikt agreements, publishers provide relevant metadata on authors and affiliations for all articles published through the agreement. These publication reports make it possible to connect each article to a single institution. A so-called paying institution. When estimating the cost of publishing Open Access outside such an agreement, one is generally confined to relying on data about corresponding author affiliations. This presents several issues and challenges. One of the most problematic is the case of multiple corresponding authors and/or multiple affiliations. This is an issue of determining which author and which institution is most likely to handle the potential payment of an APC invoice.

A widely adopted approach is using the first corresponding authork, as listed on the publication. However, this still leaves plenty of room for speculation. As shown by the plot in Figure 5.1 where the solid blue line represents the actual number of articles attributed to a Sikt consortium member institution, and the grey area above it represents the potential number of articles where at least one member institution was listed as an affiliation of the first corresponding author.

Over the last 5–10 years most of the universities and larger university colleges in Norway have had dedicated APC funds, where researchers can apply for APC funding when publishing Open Access (mainly gold journals). The Research Council of Norway also had a program called STIM-OA [93], where they reimburse institutional APC funds up to 50% of the cost (some restrictions apply).

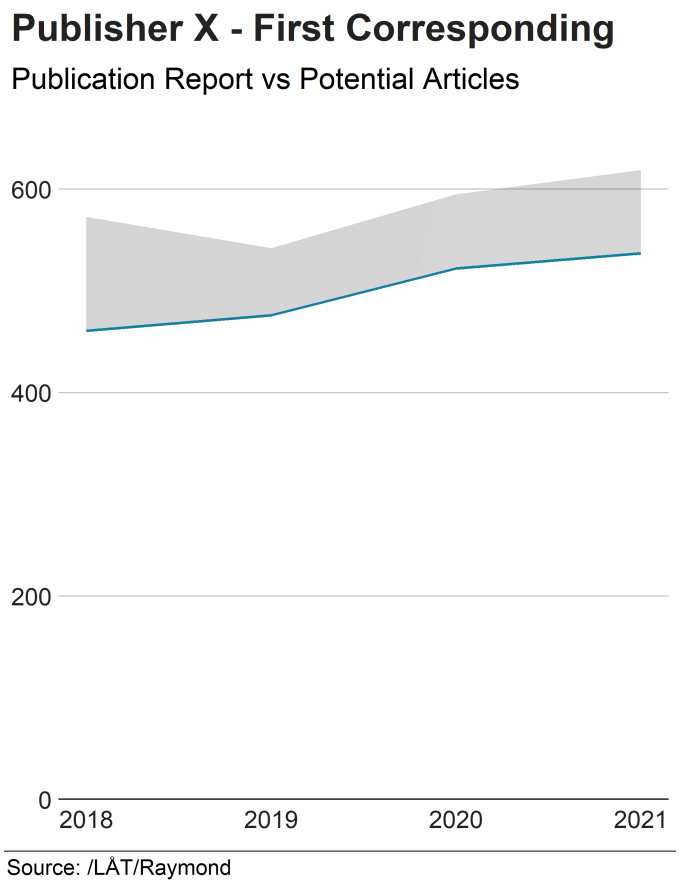

Figure 5.1: Blue line representing the number of articles where a member institution of the Sikt consortium was credited with the financial responsibility for the article publication, as per the publisher's publication report. The grey area above the blue line represents the number of articles where member institutions of the Sikt consortium was listed as affiliations of the first corresponding author.

STIM-OA provides an additional source of accurate cost data for at least some of the articles published OA by Norwegian institutions. However, the data collected by the STIM-OA program can be limited, the collection process has been very manual, which often makes the data somewhat inaccessible and cumbersome to reuse.

5.3 Assessing License Agreements

Another side of the cost of publishing coin is assessing and evaluating license agreements. At Sikt this has become an increasingly important and central part of the negotiation process. The main tool for performing such an evaluation is an inhouse developed Interaction Analysis (IA-analysis). As the name implies, it is based on different types of interactions a specific institution has with a certain journal. This institution-journal level analysis is done for every possible combination of member-institutions and journals within an agreement. The interactions currently include downloads, references (out-going citations) and published articles. Development work is already under way to include other relevant interactions in order to balance the analysis in terms of different institution profiles, e. g., curriculum lists, etc.

The underlying idea is to estimate the value of an agreement by calculating the cost of an institutions basic need for read access and the cost of publishing OA without a consortium agreement.

An example of an IA-analysis summary can be seen in Figure 5.2 and Table 5.1. It includes the current agreement cost in the first column (grey), the middle column is the cost estimated by the interaction analysis. This column consists of an estimated read cost or read value at the bottom (dark colour), and the estimated cost of publishing OA at the top (lighter coloured segment).

The analysis displays the publishing estimate based on individual journals list-price APC and the number of articles published by the institution. However, OpenAPC and other sources indicate that the list-price APC for a specific journal can differ substantially [94]. Potential explanations for such a variation can be society membership or other discounts, different prices based on article type or length, etc. This issue of obtaining or estimating accurate and realistic APC data in a specific case, again relates back to the aforementioned challenges of determining the cost of publishing open outside an agreement.

In cases where a publication agreement is already in place, the analysis is based on data from publication reports from publishers. Otherwise, this is naturally not an option.

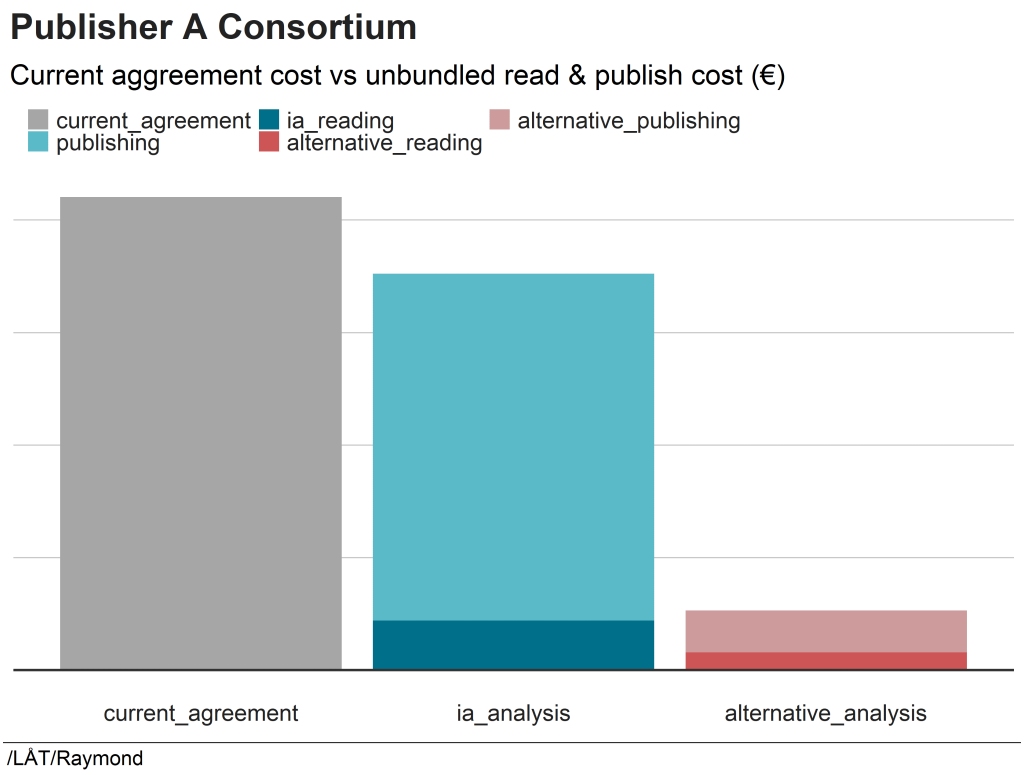

Figure 5.2: An overview of an Interaction Analysis, comparing current agreement cost with the estimated costs of basic read access and publishing open in the journals without a consortium agreement in place.

An additional, but equally important, element is being able to double check and verify the data and numbers in the publisher's reports.

With a national CRIS, in connection with the result-based funding scheme, Norway is in a unique position in terms of having control over research output, articles and affiliations of Norwegian authors. By enriching the CRIS data with corresponding authors, we have been able to predict the output volume of the consortium within an agreement with comparatively higher accuracy. Even so, there is still plenty of room for improvement.

publisher	**A**
participant	consortium
journals_total	1 348
relevant_journals	260
articles_pub_reports	714
articles_1st_corr	892
journals_no_interaction	226
journals_with_downloads	1 122
journals_with_pub_rep	257
journals_with_1st_corr_pub	268
journals_referenced	981

Table 5.1: A summary of key indicators from an Interaction Analysis, for evaluating a publish & read agreement.

References

[89] Antoine Blanchard, Diane Thierry, and Maurits van der Graaf. *Retrospective and prospective study of the evolution of APC costs and electronic subscriptions for French institutions*. Research rep. Comité pour la science ouverte, Dec. 2022. DOI: 10.52949/26 (cit. on p. 63).

[90] Norwegian Ministry of Education and Research. *Norwegian national guidelines for Open Access*. Policy. Aug. 2017. 2 pp. (cit. on p. 64).

[91] SIKT. *CRISTIN – National CRIS. Reporting instructions (NVI)*. Mar. 2019. URL: https://www.cristin.no/english/resources/reporting-instructions/ (cit. on p. 64).

[92] Nina Karlstrøm, Lars Wenaas, and Vidar Røeggen. *Overgang til åpen publisering i Norge. Status for åpen publisering og kartlegging av kostnader*. 32 pp. (cit. on p. 64).

[93] The Research Council of Norway. *Support for Open Access Publication for the Financial Year 2021. STIM-OA*. Apr. 2022. URL: https://www.forskningsradet.no/en/call-for-proposals/2022/support-open-access-publication-2021/ (cit. on p. 65).

[94] Christoph Broschinski et al. *OpenAPC/openapc-de: Couperin Additions*. Version v4.79.23-3-3. Zenodo, Mar. 2023. DOI: 10.5281/zenodo.7707499 (cit. on p. 67).

6

Analysis of Open Access publication costs at Austrian Universities

Kerstin Grossmaier-Stieg
Medical University of Graz

Christian Kaier[1]
University of Graz

DOI: 10.3204/DESY-PROC-2023-01/06

Abstract

After a short overview of Austrian Transition to Open Access Two (AT2OA2), a national project supporting Open Access, this contribution will focus on the status quo regarding Open Access Cost Monitoring at Austrian Universities as well as on approaches to improve documentation and transparency regarding publication costs. The following aspects will be addressed: capturing and displaying (Open Access) publication costs in accounting and other systems, fostering collaboration across the institution, raising awareness for publication costs at the management level, designing workflows to gather and monitor all kinds of publication costs (centrally and locally funded cost, transformative agreements etc.), as well as improving the reporting of (Open Access) publication costs in statistics.

[1] https://orcid.org/0000-0002-8750-6666

6.1 Introduction

Austrian Transition to Open Access Two (AT2OA2, 2021–2024) [95] is a project funded by the Austrian Federal Ministry of Education, Science and Research with the purpose of advancing the transformation from closed access to Open Access. Among other things, the preceding project AT2OA (2017–2020) supported the establishment of Open Access publication funds at Austrian universities. While formulating model funding conditions [96] within the AT2OA project it became obvious that there would be ineligible publishing costs that would continue to be financed through other sources and therefore be "invisible" from a library perspective. At about the same time, concepts such as APCs in the wild [97] or Total Cost of Publication [78] gained attention, followed by the vision of an information budget [12] as well as to the establishment of new service providers that offer solutions for central invoicing workflows and improved cost transparency.

During 2017/18 some of the universities involved in AT2OA started investigating publication costs across the institution with the aim of overseeing and monitoring these expenditures. Within AT2OA a "Recommendation for booking Open Access publication costs" was published in July 2020 [98].

At that time, funding for AT2OA2 had already been approved. As in the preceding project, all public universities in Austria, the Research Institute of Molecular Pathology and the Institute of Science and Technology Austria are project partners in AT2OA2, which facilitates joint action by the autonomous institutions with regard to Open Access. AT2OA2 includes the following five subprojects:

- More Transformative Agreements for Austria (SP1)
- Austrian Datahub for Open Access Negotiations and Monitoring (SP2)
- Analysis of Publication Costs at Austrian Universities (SP3)
- Predatory Publishing (SP4)
- Visibility of Open Access Publications (SP5)

6.1.1 Subproject 1 — More Transformative Agreements for Austria

The goal of SP1 is to increase the number of Open Access publishing agreements of Austrian institutions with scholarly publishers. The predecessor project AT2OA

already contributed substantially to the funding of Open Access publishing agreements with Springer, Wiley and Elsevier. The measures taken from AT2OA are continued in the current project with the aim of concluding more such agreements. The project team has defined new guidelines for Open Access publishing agreements [99] relating to Open Access in general, funding mechanisms, invoicing, author identification, article-related information and publishing mechanisms for articles, reporting and statistics. According to the guidelines, an agreement with a scholarly publisher must (among other criteria) include a commitment to Open Access, transparency of publisher services and costs, and concepts for a transformation towards Open Access as well as automated workflows.

6.1.2 Subproject 2 — Austrian Datahub for Open Access Negotiations and Monitoring

Based on the recommendations developed in AT2OA, a national data hub is being established in Austria. The Austrian Datahub for Open Access Negotiations and Monitoring will process publication data from various sources in order to be able to use them for national Open Access monitoring to support negotiations with scholarly publishers. Publication metadata is provided in high quality in the current research information systems (CRIS) that are set up at the institutional level. However, information on Open Access publication conditions, such as APCs, access and licensing and embargo dates, is missing in most cases. In the data hub information from CRIS is combined with information from external database providers, service providers and publishers. All participating institutions will be able to benefit from the subsequent use of the enriched data. Upon completion of the project, the data hub will have been established and transferred into regular operation. The project group intends to collaborate with international projects and similar activities.

6.1.3 Subproject 3 — Survey and analysis of publication costs at Austrian universities

In SP3, analyses of publication costs at the project partner institutions are carried out and practical guidelines are developed on how university libraries, accounting & controlling departments, quality management and research information systems can be better coordinated in order to be able to clearly identify and allocate publication costs. The project group is working on better cooperation, knowledge

exchange and coordination of the units involved within the organizations in order to be able to clearly identify and allocate publication costs. SP3 is also working on how to display publication costs in statistics and how to best track and deal with third-party publication costs.

6.1.4 Subproject 4 — Predatory Publishing

The main goal of SP4 is to raise awareness about predatory publishing and other predatory phenomena, build an information infrastructure for university stakeholders (such as university management, library management, academic staff and students) and train staff members at the partner institutions for individual consultations of researchers. While there have been institutional activities up and running for quite some time, predatory publishing and predatory conferences are examined within SP4 in a national context for the first time. The project group conducted a survey on the services and support offered around predatory publishing at the project partners' institutions. Members of the project group also actively contacted Austrian scholars who (unknowingly) appear on editorial boards of suspicious journals. As predatory phenomena are not just a national problem the group also focuses on connecting and exchanging knowledge with other communities outside the country. Insights will be presented at various conferences, and a blog was initiated that addresses key questions about predatory publishing and predatory conferences [100].

6.1.5 Subproject 5 – Visibility of Open Access Publications

Subproject 5 deals with the central question of whether Open Access promotes not only better availability of scholarly literature but also an increase in visibility. Insights are gained based on the data collection of the previous project AT2OA (extended by current volumes) with the help of alternative metrics (Altmetrics). Altmetric [101] (part of the Digital Science group) accompanies the project with a Scientific Advisor. The project team intends to look into whether the Open Access status correlates with increased online attention in Altmetric, whether there are differences between individual expressions of Open Access in relation to Altmetrics, if disciplinary and temporal differences can be identified, whether a normalization of Altmetrics (comparable to citation metrics) succeeds in increasing interpretation and meaningfulness for the insights gained and what other means could be used (apart from Altmetrics) to measure visibility. At the end of SP5 the project team will have published an final report on the project, the

publication datasets for project members will be enriched with Altmetrics data, there will be lessons learned (probably even workflows and best practices) for data enrichment, and the project group will have made recommendations for responsible use of Altmetrics.

6.2 Challenges

While publication funds to cover costs for Open Access articles and books have been and continue to be established at many higher education and research institutions [102], APCs in the wild have increasingly moved into the focus [12, 79, 98, 103–105]. These costs, covered decentrally (on the level of different departments, individual researchers etc.), often make up for a considerable proportion of all Open Access payments across an institution. They may result from authors not being aware of funding opportunities such as transformative agreements and institutional publication funds; from authors, publications or charges not eligible for centralized funding as well as from innovative forms of publication funding including membership models. Print subsidies, color charges, pages charges and other types of costs not related to Open Access contribute to a fuller picture of publication costs. Expenditures on Open Access deals with publishers are also part of the idea of a Total Cost of Publication [78], but it is difficult to separate them into fees for publishing and fees for reading. In consequence, it is very challenging to attribute a concrete amount of money to each single Open Access publication. The concept of an information budget [84] aims at reconciling these different kinds of costs by calculating a total amount of expenditure for information supply. However, such an information budget is extremely hard to delimit and still needs to be defined in detail. As a first step, AT2OA2 engages in the field of publication costs, with an emphasis on Open Access costs.

Considering the difficulties libraries – as one of the key players in the transition to Open Access – face in determining what should be counted as publication costs and in assigning costs to publications and vice versa, it is not surprising that other organizational units such as accounting departments find it hard to capture and correctly assign such costs. One of the main tasks the project partners have set for themselves is to deepen the exchange of information across academic institutions and to inform administrative departments on how to identify and handle publication costs. Obviously this task requires much effort and engagement from a wide range of stakeholders and is still to be completed.

6.3 Project aims

According to the AT2OA2 project proposal submitted in September 2019, SP3 is committed to

- supporting institutional reporting systems by making the data reusable for statistics, annual reports and the like
- increasing knowledge regarding Open Access publication costs: who pays for what?
- creating synergy effects in the relevant administrative processes and workflows

and addresses the following questions:

- How much money is paid at an institution for publishing?
- How much do publishers charge for publishing?
- Which funder or institution has covered the publication costs?

A particularly important objective of SP3 is to help create a reliable and comprehensive data basis for publication costs, especially for Open Access publications. As a consequence, (ideally) all publication costs covered by institutions on central and decentral levels should become visible and transparent. Detailed and complete data on what institutions and their researchers are charged by publishers will improve the position of institutions in negotiations. Beyond that, SP3 would like to contribute to a better display of (Open Access) publication costs in statistics such as the Austrian Library Statistics [106] and all sorts of reporting tools.

All these endeavors include and require fostering cooperation among all involved units within the institution. The approach to work together with other administrative as well as research departments characterizes the collaborative nature of SP3 and distinguishes it from other initiatives that aim to improve structures and systems within libraries. Enhanced cooperation should in turn create synergy effects within the administrative processes, further improving cost efficiency in the transition to Open Access.

6.4 Activities to date

Between May 2021 and November 2022, eight meetings of SP3 were held to establish a common understanding of the problem, identify potential solutions, define deliverables in detail and start working on concrete tasks.

As a starting point SP3 conducted a survey in summer 2021 among the participating institutions to collect information on the status quo regarding cost monitoring and transparency. The next subchapter will elaborate on the most important outcomes of this survey.

In early 2022 SP3 invited the project partners to collect the cost data available for their institutions and to share their knowledge and experiences. We found that there is a wide range of ambition and understanding at the various institutions for the relevance as well as for the challenges of our topic.

In June 2022 SP3 organized an international workshop on cost monitoring with Irene Barbers (Forschungszentrum Jülich)[2], Lisa Lovén (National Library of Sweden), Susanne Luger (University of Linz), Timo Vilén (National Library of Finland)[3] and Mathew Willmott (California Digital Library) sharing their insights, experiences and approaches on cost monitoring.

Thanks to an invitation to the openCost workshop in October 2022, SP3 gained insight into ongoing efforts concerning cost capture, documentation and monitoring in Germany and beyond. While the openCost project complements the SP3 approach of bringing together people involved in the topic with their emphasis on metadata and technical interfaces, it was also particularly interesting to learn about the Transform2Open project[4]. During the presentations it turned out that we share a substantial number of approaches and ideas on the topic, and we are hoping to continue the inspiring and fruitful exchange initiated in Hamburg.

6.5 Survey on Open Access Cost Monitoring

From July to September 2021, SP3 conducted an internal survey with 24 questions in order to get an overview of how publication costs are dealt with at the partner institutions. Representatives of 16 Austrian research institutions answered the survey.

[2] see also her contribution in chapter 10
[3] see also his contribution in chapter 7
[4] see also chapter 8

Some of the results:

- Two thirds of the institutions either use SAP (dedicated cost types Open Access or Publications) or Excel lists in order to collect centrally as well as decentrally incurred publication costs. Two institutions managed Open Access costs via a central office. One third of the participants indicated that their institutions did not collect publication costs.
- Half of the participants indicated that their institutions had central monitoring for publication costs.
- Two thirds of all institutions had already set up one or more SAP cost types for Open Access publication costs.
- At most institutions, lump-sum costs from Open Access agreements were not allocated to individual publications (one participant noted that "a coordinated approach would be desirable").
- A *wish list* of what should have improved after the completion of SP3 includes
  - uniform standards/recommendations/guidelines/proposals for transparency of publication costs
  - suggestions for displaying (Open Access) publication costs in SAP, e. g. via dedicated cost types or improved accounting texts
  - overview and regular monitoring of all publication costs at the institution
  - increased professionalism and know-how in data collection & monitoring
  - automation of workflows wherever possible
- Data in connection with publications from third-party funded projects is mostly not available to libraries (only two participants indicated they had access) but only to accounting departments and/or research support offices.
- With regard to these costs, the survey participants would hope for
  - improved query via their CRIS system
  - central query options in SAP
  - full list of publication costs from global budget and third-party funded projects

The survey results show that the status quo concerning publication cost transparency and workflows differs a lot among the project partners due to individual, historically grown structures and responsibilities. The results also indicate that awareness for the issue of Open Access cost monitoring is still to be raised at the management level.

A wide range of answers was given to the question of how AT2OA2 could improve the situation. This feedback was then used to revise and further develop the project deliverables.

6.6 SP3 Deliverables and status quo

SP3 has defined five topics as the main priorities to work on and deliver results by the end of 2024:

6.6.1 SP 3.1 Setting up SAP cost types

Partner institutions within AT2OA2 have reached a common understanding to use SAP as the backbone and primary resource for cost data analysis. A project group within the preceding project AT2OA already published recommendations for setting up cost types in SAP [98]. The working group in AT2OA2 has built on these recommendations and has revised them in order to meet current needs and emphasize advantages of an improved cost monitoring. The document includes arguments for the establishment of Open Access cost types, taking into account the varied Open Access business models. The recommendations address rectorates (university management), controlling and accounting departments as well as libraries. The current recommendations, which include the challenge of monitoring third-party publication costs and also put a strong focus on the management aspect (the need to control publication costs) are to be finalized by the end of 2022 and will be published on the AT2OA2 website and on Zenodo.

6.6.2 SP 3.2 Developing training materials

The working group aims at defining guidelines and arguments for a standardization of accounting texts on the basis of existing documents at individual institutions and the recommendations published within SP 3.1 (subsection 6.6.1). This is meant to support colleagues at accounting, finance and controlling departments, but also decentralized units such as institutes, chairs, centers and clinics in processing and allocating Open Access invoices correctly. The working group currently discusses whether it is reasonable and desirable to approach publishers to deliver

uniform and standardized invoices for Open Access publication costs. The training material should be finished by summer 2023.

6.6.3 SP 3.3 Modeling workflow descriptions

The project team is developing descriptions of prototypical workflows and flowcharts for collecting and analyzing publication costs for Hybrid Open Access, Gold Open Access as well as other publication costs covered both by institutional funds and third-party funding. Different aspects such as the diversity of business models (publishing agreements, individual funding, deposits, caps, etc.), extra charges such as color charges, page charges, submission fees, etc. and crowdfunding models will be taken into consideration. The aim of the working group is to improve documentation and knowledge management, but also to support colleagues who are new to the topic of publication cost monitoring.

6.6.4 SP 3.4 Preparing costs for statistics and reporting

There are no uniform and comprehensive standards yet on how to report (Open Access) publication costs for library statistics and institutional reporting. The project group has already started working on a standardized representation of publication costs for the delivery of key indicators. As a first step, a proposal on how to improve the display of expenditures for Open Access in the Austrian and German library statistics will be completed in early 2023. The key figures to be standardized will include library expenditures for all forms of media, electronic media and Open Access publications. The proposal will suggest to differentiate between APC/BPC payments and funding for Open Access infrastructures, read and publish agreements (including transformative agreements) and Open Access costs within an institution's own research infrastructure. The project group intends to collaborate with experts in Germany, as the library statistics for both countries are based on the same scheme.

6.6.5 SP 3.5 Handling of publication costs from third-party funded projects

Publication costs from third-party funding are a particular challenge and therefore require special attention and handling. The number of funders can be manifold at institutions, and so are the different funding policies. The Austrian Science Fund will outsource the complex process of reviewing publications and journals for compliance with the FWF's Open Access Policy as well as the reimbursement

of publication costs to higher education and research institutions in Austria as of 2024. SP3 will therefore in particular look into the aspect of how third-party funding in connection with publication costs can be dealt with reasonably and efficiently. The project group has already contacted the Austrian Science Fund to receive detailed information on the parameters of this outsourcing, including the topic of cost monitoring. It will also inform and collaborate with institutional management, colleagues from accounting, finance and controlling departments as well as colleagues who support researchers in externally funded projects. By the end of the project, a concept shall be provided on how to best handle publication costs from third-party funders.

6.7 Conclusion

Publication costs are rapidly gaining importance at higher education and research institutions. These costs are complex and dynamic, and they require know-how and new forms of collaboration across stakeholders on an institutional, national and international level.

This is why – beyond working on the project deliverables mentioned above as well as offering a forum for sharing experiences, approaches and best practices mainly among librarians – SP3 explicitly attempts to go beyond its own community.

Networking is central at this stage, just as working on guidelines, standards, schemas and workflows is. SP3 is looking forward to further advancing the transparency of publication costs locally as well as together with the growing number of initiatives dealing with this important aspect of the transition to Open Access.

6.8 Acknowledgements

The authors would like to thank all members of SP3, our colleagues from the subprojects, Brigitte Kromp and Maria Seissl as AT2OA2 project leaders, and the AT2OA2 office for the excellent collaboration and the continuous support.

References

[12] Heinz Pampel. *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren*

Kosten. de. Arbeitspapier. 2019. DOI: 10.2312/OS.HELMHOLTZ.006 (cit. on pp. 6, 17, 47, 72, 75).

[78] Stephen Pinfield, Jennifer Salter, and Peter A. Bath. "The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions". In: *Journal of the Association for Information Science and Technology* 67.7 (2016), pp. 1751–1766. DOI: 10.1002/asi.23446 (cit. on pp. 47, 72, 75, 87).

[79] Bernhard Mittermaier and Irene Barbers. "Zeitschriftenmonitoring". In: *Bibliotheken. Wegweiser in die Zukunft. : Projekte und Beispiele*. Ed. by Erda Lapp et al. Berlin: Simon Verlag für Bibliothekswissen, 2021, pp. 346–359. DOI: http://hdl.handle.net/2128/22746 (cit. on pp. 47, 75).

[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 (cit. on pp. 47, 75, 94, 95, 121).

[95] AT2OA². *Austrian transition to open access 2*. URL: https://www.at2oa.at/at2oa2_home.html (visited on 2022-12-06) (cit. on p. 72).

[96] Christof Capellaro et al. *Open-Access-Publikationsfonds. Einrichtung und Förderbedingungen*. de. 2019. DOI: 10.5281/ZENODO.2653724 (cit. on p. 72).

[97] Jessica Monaghan, Mithu Lucraft, and Katie Allin. *'APCs in the Wild'. Could Increased Monitoring and Consolidation of Funding Accelerate the Transition to Open Access?* White paper. 2020. DOI: 10.6084/M9.FIGSHARE.11988123.V4 (cit. on p. 72).

[98] AT2OA. *Empfehlung zur Buchung von Open-Access-Publikationskosten*. de. 2020. DOI: 10.5281/ZENODO.3945952 (cit. on pp. 72, 75, 79).

[99] AT2OA². *Überarbeitete Leitlinien für Verlagsverträge für die OA-Transformation im Rahmen von AT2OA2*. URL: https://www.at2oa.at/docs/at2oa2/Ueberarbeitete%20Leitlinien%20fuer%20Verlagsvertraege%20fuer%20die%20OA-Transformation%20im%20Rahmen%20von%20AT2OA2.pdf (visited on 2022-12-06) (cit. on p. 73).

[100] AT2OA². *Scholarly communications in transition: predatory and other phenomena in academia*. URL: https://in-transition.at/ (visited on 2022-12-06) (cit. on p. 74).

[101] Digital Science & Research Solutions Ltd. *Altmetric*. URL: https://www.altmetric.com/ (visited on 2022-12-06) (cit. on p. 74).

[102] Heinz Pampel. *Open Access an wissenschaftlichen Einrichtungen in Deutschland. Ergebnisse einer Erhebung im Jahr 2018*. de. Bericht. 2019. DOI: 10.2312/OS.HELMHOLTZ.005 (cit. on p. 75).

[103] Lisa Lovén. "Monitoring open access publishing costs at Stockholm University". In: *Insights the UKSG journal* 32 (2019). DOI: 10.1629/uksg.451 (cit. on p. 75).

[104] Claudia Frick and Christian Kaier. "Publikationskosten für Zeitschriftenartikel abseits von Open-Access-Publikationsfonds – Lost in Transformation?" de. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 7.2 (2020), pp. 1–15. DOI: 10.5282/O-BIB/5586 (cit. on p. 75).

[105] Andre Bruns and Niels Taubert. "Investigating the Blind Spot of a Monitoring System for Article Processing Charges". In: *Publications* 9.3 (Sept. 2021), p. 41. DOI: 10.3390/publications9030041 (cit. on p. 75).

[106] *Bibliotheksstatistik.de*. URL: https://www.bibliotheksstatistik.de/ (visited on 2022-12-06) (cit. on p. 76).

7

Cost Monitoring in Finland

Timo Vilén[1]
The National Library of Finland

DOI: 10.3204/DESY-PROC-2023-01/07

Abstract

In my presentation, I will look at the current state of cost monitoring in Finland, drawing on my experience as the leader of FinELib's APC project and the chair of an expert group charged with reviewing the total costs of OA to Finland. I will describe our approach including our attempts to harness the Finnish VIRTA Publication Information Service for the purposes of cost monitoring, not forgetting some of the challenges encountered along the way. Some of the questions I will address are: What do we know (or think we know) already and what other aspects should we consider to get a better picture of the various costs associated with scholarly publishing?

7.1 Monitoring APCs

In Finland, the discussion on the need for greater transparency and monitoring of APCs started in earnest around 2017–2018, prompting FinELib (the Finnish consortium for universities, research institutions and universities of applied sciences) to initiate a project referred to unimaginatively as FinELib's APC Project (which I coordinated) [107]. Funded by the Finnish Ministry of Education and Culture

[1] https://orcid.org/0000-0003-0324-2489

and modeled, to some extent, on a similar project conducted by our sister consortium Bibsam [108], the project (2019–2020) aimed at exploring the motivations and challenges associated with the monitoring of APCs among our consortium members.

The project also sought to estimate the total APC expenditure of Finnish institutions, while at the same time offering recommendations on what measures should be taken to facilitate the tracking of APCs on a national level. These included, among other things, recommendations to set up specific accounts and accounting codes to enable tracking as well as to take advantage of the opportunities offered by Current Research Information Systems (CRIS) and equivalent for a more streamlined monitoring. Crucially, the report also introduced common minimum standards for data to ensure a consistent, interoperable approach across Finland [107].

As one of the aims of our project had been to explore various ways of sharing Finnish APC data through OpenAPC, it was only natural that the adopted standards and minimum requirements had to be aligned with those proposed by the OpenAPC initiative [109]. Institutions participating in the project also submitted their data to OpenAPC and, at the time of writing this paper, almost all our major universities – and several universities of applied sciences and RIs – share their APC data with the OpenAPC initiative [110] either directly or via Finland's National VIRTA Publication Information Service [111]. And, I would argue, they do this not just out of solidarity or because they are asked to but because they, too, agree on the importance of cost transparency in creating more informed buyers.

7.2 Towards the Total Costs of Publication

With the monitoring of APCs well underway, next up on our national checklist was reviewing the total costs of publication in Finland. This had been one of the responsibilities assigned to the National Library (read FinElib) in Finland's national policy for open access to scholarly publications [112], and was also acknowledged by our APC project which recommended that a task group under the leadership of the National library be set up to review the total costs of OA to Finnish institutions. In addition, the report proposed that the group would develop a model for monitoring the total costs of publication on an ongoing basis [107].

This we duly did, and in early 2022 a task group consisting of experts from libraries and key stakeholders set out to consider what kind of approach would be the most appropriate for our needs. The monitoring set up earlier by Bibsam again

served as an inspiring example [113], while also prompting us to ask what exactly we mean by the total costs. To put it another way: what are the components that together make up what we refer to as the *total* or *full costs* of scholarly publication as seen from the perspective of institutions? Which costs should (or could) we include and not include in our approach?

There are, of course, a number of tangible and easily attributable costs that institutions already monitor and that should obviously be taken into consideration when talking about the total cost of publication, including APCs. Self-evidently, subscriptions placed at the consortial level had to be part of the equation and, indeed, have been made publicly available in our annual reports [114], while *all* license and subscriptions fees paid between 2010-2017 by Finnish institutions to publishers and vendors (be via FinELib or directly) were collected and shared openly by the Finnish Open Science and Research Initiate ATT, an exercise terminated in 2017 due to lack of funds [115].

Moreover, it can be easily argued that the fees paid to the infrastructures we and our current OA and OS ecosystem rely on so heavily (DOAJ, DSpace etc.) ought to be included, which, of course, is not to say that they can necessarily be easily pulled from an institution's accounting system.

But what about more intangible administrative costs borne by institutions such as staff time spent on processing APC payments, helping authors to navigate an increasingly complex OA environment, negotiating and managing OA agreements [78], or enabling "green" OA (e. g. depositing copies of articles into repositories for storage and access)? These are real, legitimate costs which we often struggle to identify and which, indeed, resist quantification but nevertheless costs that ought to be recognized as part of the total costs of publication.

Further complicating the picture is the fact that monitoring itself comes at a cost. This of course begs the question: even if tracking certain costs would be possible, does it follow that they *should* be tracked? The answer, of course, depends on many things, including what we want to achieve, the value we attach to transparency, and, obviously, the resources we have available.

This is not to argue that we should prefer opacity to transparency or that there is currently enough transparency about costs. Rather, this is to say that while greater transparency is often desirable, there are situations where the benefits of greater transparency are outweighed by the costs associated with achieving it [116].

7.3 Defining Costs

To bring some clarity into the jungle of total costs of publication, we might do well to distinguish between three different definitions of costs inspired by the approach by Maron et al. [117]:

- Basic costs: Includes "out-of-pocket" (direct) costs such as APCs, BPCs, subscription fees etc. which are relatively easy to track and, more often than not, can be traced to a particular publication. Accounting for Basic costs provides some level of transparency which, for want of a better expression, can be termed Basic transparency.

- Full costs: Covers Basic costs and (mostly indirect) administrative costs related to the management and implementation of OA (e.g monitoring, administration, communication, and enabling green OA). The level of transparency achieved is, of course, higher, so much so that the term "Full transparency" might be warranted as long as one takes care not to omit quotations marks.

- Full costs Plus: Includes not only Full costs but also incorporates "in-kind" contributions and unpaid (voluntary) work associated, for instance, with the running of Diamond OA journals, the preferred business model for the overwhelming majority of Finnish OA journals. Going for this approach would provide even more transparency (let's call this degree of transparency "Full transparency Plus"), though tracking the said costs would, of course, prove difficult, if not impossible.

To use a somewhat clichéd analogy, Basic costs are the low-hanging fruits in our quest to monitor the total costs: easy to pick, relatively speaking, and often ripe enough to eat. Hanging on the upper branches of our cost tree are Full costs. These "fruits" are harder to pick but they, too, are within our reach, provided that we are willing to make an effort – and provided that we have proper ladders or a telescopic fruit picker at hand.

Finally, Full costs Plus are the fruits that hang too high on our tree to pick by normal means. However, by observing the tree, we can estimate the number of such fruits, while by giving our tree a proper shake, we might even be able to draw some conclusions as to their quality based on the (presumably ripe) fruits that fall.

7.4 Our Approach?

This was, in essence, the approach taken by our task group, though it must be added that I only came up with the above distinction and analogy when writing this paper. Either way, the approach put forward by our group [118] shared many similarities with the framework used by Bibsam ever since 2017 [113]. In addition to the most obvious low-hanging fruits, we, too, asked our institutions to provide data on costs related to the implementation and management of OA (estimated staff time = person years). In keeping with the Bibsam approach, we also included fees paid for printed materials to provide some insights into the possible reallocation of existing resources.

At the same time, we wanted to go a step further by including the costs of OA infrastructure as well as those of Green OA, in particular, the extent to which depositing versions of manuscripts in publicly accessible repositories represents costs to institutions. In addition, and at the risk of trying the patience of our future respondents, we decided to include the fees paid to Finnish scholarly journals, be it subscriptions or APCs [118]. As already mentioned, a sizeable majority of the Finnish OA journals have embraced the Diamond OA model, and by incorporating these costs we were hoping to contribute to the ongoing and, at times, heated debate on their funding basis which, in its current state, is far from being equitable and sustainable.

Data for the monitoring came from a variety of sources:

1. From a survey conducted by the Open Science and Research Coordination in Finland in 5–6/2022 as part of the broader monitoring of the state of open science and research in Finland [119]. The survey was sent to all Finnish universities, universities of applied science, and RIs, and despite some demanding questions, attracted an overall response rate of 100 per cent among universities and nearly 100 per cent among other institutions;

2. From the FinELib office (or to be more precise, from our very own Halti system), and;

3. From the OpenAPC through which, as already stated, a number of Finnish institutions have been sharing their APC cost data [110].

Once gathered, the data was compiled and cleaned by CSC, the Finnish IT Centre for Science, while the results, including interactive charts, and data were published on research.fi [118] and Fairdata.fi [120], both platforms provided by

the Finnish Ministry of Education and Culture [118] In future, the monitoring will be conducted biennially to reduce the burden on institutions – unless, of course, we as a community agree otherwise – while possible future changes to the model have to be discussed with the key stakeholders. These could include, for instance, asking additional questions on other publication charges (e. g. page and colour charges), increasing automation, or updating how we display our data.

7.5 Lessons Learned

So, what did we find out about the total costs of publication in Finland? This was not known at the time of our workshop, and even though the results of our monitoring have since been published, I shall not discuss them here but, rather, invite the reader to engage with them on research.fi, and while doing so, to keep in mind that what they are seeing is an interim version and still under development [118]. (It goes without saying that we would welcome any comments.)

Instead, I would like to end by reiterating some of the lessons learned from our exercise. Firstly, we need to think more seriously about what we are including and not including when the issue of total costs of publication is raised.

Secondly, while data is essential, in every monitoring exercise there is an apparent danger of becoming too obsessed with data gathering. We should, therefore, keep asking ourselves and our poor colleagues why monitoring matters, what roles it is supposed to serve, and, crucially, what it is that we don't know but should know to be able to make informed and effective decisions. To put it more simply, instead of finding questions for data, we should find data for questions. This approach has been called *decision-driven data analytics* by de Langhe and Puntoni in their recent article [121], and I fully agree with their conclusions.

Finally, and as already indicated, reducing information asymmetries tends to have direct and indirect costs. Thus, for all the talk about promoting transparency, we should not lose sight of the fact that maximal transparency, lofty as it may sound, may not be the same as optimal transparency, suggesting that the notion of transparency, too, merits consideration.

References

[78] Stephen Pinfield, Jennifer Salter, and Peter A. Bath. "The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions". In: *Journal of the Association for Information Science and Technology* 67.7 (2016), pp. 1751–1766. DOI: 10.1002/asi.23446 (cit. on pp. 47, 72, 75, 87).

[107] Timo Vilén et al. "Kirjoittajamaksut ja niiden seuranta: havaintoja ja kehitysehdotuksia. FinELibin Open APC -projektin loppuraportti". In: *Kansalliskirjaston raportteja ja selvityksiä* 28 (2021) (cit. on pp. 85, 86).

[108] National Library of Sweden. *Open APC Sweden*. URL: https://www.kb.se/samverkan-och-utveckling/oppen-tillgang-och-bibsamkonsortiet/open-apc-sweden.html (visited on 2023-01-30) (cit. on p. 86).

[109] OpenAPC. *OpenAPC data schemas*. URL: https://github.com/OpenAPC/openapc-de/wiki/schema#openapc-data-set (visited on 2023-01-30) (cit. on p. 86).

[110] OpenAPC. *OpenAPC Finland*. URL: https://treemaps.openapc.net/apcdata/openapc/#institution/country=FIN (visited on 2023-01-30) (cit. on pp. 86, 89).

[111] *VIRTA Publication Information Service*. URL: https://www.csc.fi/en/-/virta-publication-information-service (visited on 2023-01-30) (cit. on p. 86).

[112] Open Science Coordination in Finland and Federation of Finnish Learned Societies. *Open access to scholarly publications. National Policy and executive plan by the research community in Finland for 2020 – 2025 (1)*. 2nd edition. Vol. 3. Responsible Research Series. Nov. 2019. DOI: 10.23847/isbn.9789525995343 (cit. on p. 86).

[113] National Library of Sweden. *Utgifter för vetenskaplig publicering*. URL: https://www.kb.se/samverkan-och-utveckling/oppen-tillgang-och-bibsamkonsortiet/utgifter-for-vetenskaplig-publicering.html (visited on 2023-01-30) (cit. on pp. 87, 89).

[114] *FinELib-konsortion vuosi 2021*. 2021. URL: https://finelib.fi/wp/wp-content/uploads/2022/06/FinELibin_vuosikertomus_2021.pdf (visited on 2023-01-30) (cit. on p. 87).

[115] Ministry of Education and Culture of Finland and Open Science and Research Initiative Finland. *Academic publisher costs in Finland 2010-2016*. Ministry of Education and Culture of Finland and its Open Science and Research Initiative 2014-2017 (cit. on p. 87).

[116] J. Forssbæck and L. Oxelheim. *The Multi-Faceted Concept of Transparency*. IFN Working Paper 1013. Research Institute of Industrial Economics (IFN), Stockholm, 2014. DOI: 10.2139/ssrn.2405951 (cit. on p. 87).

[117] N.L. Maron et al. *The Costs of Publishing Monographs: Toward a Transparent Methodology*. Research rep. Ithaka S+R., Feb. 2016. DOI: 10.18665/sr.276785 (cit. on p. 88).

[118] Ministry of Education and Culture of Finland. *Kustannukset*. URL: https://www.tiedejatutkimus.fi/fi/science-innovation-policy/open-science-and-research-indicators/indicators_content_4 (visited on 2023-01-30) (cit. on pp. 89, 90).

[119] Secretariat for the National Open Science and Research Coordination Federation of Finnish Learned Societies. *Open Science Monitoring Begins!* May 2022. URL: https://avointiede.fi/en/news/open-science-monitoring-begins (visited on 2023-01-30) (cit. on p. 89).

[120] *Fairdata.fi*. URL: https://www.fairdata.fi/en/ (visited on 2023-01-30) (cit. on p. 89).

[121] Bart de Langhe and Stefano Puntoni. "Leading with decision-driven data analytics". en. In: *MIT Sloan Management Review* 62.3 (2021) (cit. on p. 90).

8

Transform2Open — Cost monitoring, criteria, competencies, and processes of the Open Access transformation

Heinz Pampel[1]
Helmholtz Association, Helmholtz Open Science Office & Humboldt-Universität zu Berlin, Berlin School of Library and Information Science

Tobias Höhnow[2]
University of Potsdam, University Library

Lea Maria Ferguson[3]
Helmholtz Association, Helmholtz Open Science Office

Bernhard Mittermaier[4]
Forschungszentrum Jülich, Central Library

Irene Barbers[5]
Forschungszentrum Jülich, Central Library

DOI: 10.3204/DESY-PROC-2023-01/08

[1] https://orcid.org/0000-0003-3334-2771
[2] https://orcid.org/0000-0002-3483-6223
[3] https://orcid.org/0000-0002-7060-3670
[4] https://orcid.org/0000-0002-3412-6168
[5] https://orcid.org/0000-0003-2011-7444

Abstract

The Transform2Open project addresses the development of budgets, criteria, competencies, and related processes in research-performing organizations around the financial dimensions of the Open Access (OA) transformation. Transform2Open organizes dialogue forums and at the same time develops recommendations for strategies, concepts, and measures to shape the OA transformation at universities and non-university research institutions. Topics of the project include: Improvement and development of cost control methods; promoting the interaction of library budgets, third-party funding, and other financial resources in research institutions in order to create an information budget; developing recommendations for contracts with publishers; optimization of workflows for handling publications and related metadata and invoices; promoting transparency in the financial framework for the transformation towards OA; and determining the competence profiles of professionals involved in the transformation towards OA in research institutions. Transform2Open has been approved by the German Research Foundation (DFG) in 2022 and will formally commence in 2023. Partners of the project are the Central Library of Forschungszentrum Jülich, the Helmholtz Association's Open Science Office, and the University Library at the University of Potsdam. This article provides an overview of the project.

8.1 Introduction

The Open Access (OA) transformation must be shaped locally by universities and other research-performing organizations (RPO) and be embedded in national and international initiatives. Institutions have to implement operational OA activities based on policies and strategic decisions, and ensure that the OA activities are monitored. Addressing the economic dimension and its impact on RPOs is of central importance in the current phase of transformation.

While OA is becoming increasingly widespread, there is also a fragmentation of implementation measures that would potentially have for a more substantial and sustainable impact if bundled together. The results of the Options4OA project show that the handling of article processing charges (APCs) and the associated monitoring of publication costs in Germany is still at its beginning [84].

To foster the development of OA in Germany in a coordinated manner, dialogue forums are needed to discuss and develop best practices for shaping the transformation of scholarly publishing from closed to open. The need for such dialogue forums to address the financial aspects of the OA transformation is currently

evident in the discussion of the concept of information budgets in Germany [70, 84] Challenges for RPOs are, for example, the implementation of national agreements with publishers at the local level and the consideration of international developments (e. g., Plan S). Considering the continuous increase of the OA share in German publishing, a joint and coordinated approach paves the way for RPOs to successfully deal with the challenges associated with the OA transformation.

8.2 Project

Transform2Open supports transformation activities at research institutions in Germany through the following actions [49]:

1. Improving and developing cost monitoring methods;
2. promoting the interplay of library budgets, third-party funding, and other financial resources at research institutions to create overarching information budgets;
3. developing international criteria for contracts with commercial publication service providers;
4. optimizing workflows around the handling of publications and associated metadata and invoices;
5. promoting transparency around the financial framework of the OA transformation and identifying organizational structures;
6. determining competence profiles for professionals involved in OA transformation at research institutions.

Transform2Open organizes dialogue forums and develops recommendations for strategies, concepts, and measures to shape the OA transformation at universities and non-university research institutions.

Transform2Open enables the networking between actors in RPOs and the staff active in the context of information infrastructures, e. g., research libraries, regarding model development and standardization. The project further supports the engagement with the economic dimension of the OA transformation in Germany. Additionally, it promotes the transfer of knowledge within the international context.

The Transform2Open project ensures the successful interaction of various transformative efforts with DEAL [122, 123] openCost [69], and other initiatives and projects in Germany and internationally.

Partners of the Transform2Open project are the Central Library of Forschungszentrum Jülich, the Helmholtz Open Science Office, and the Potsdam University Library.

8.3 Work program

To support transformative activities at research institutions in Germany, Transform2Open operates in the aforementioned six areas (see above) and implements these via the following work packages (WP):

WP 1 **Cost monitoring**

This WP deals with strategies for monitoring publishing costs at scientific institutions. The goals of WP 1 are to document the current status of cost monitoring and to develop proposals for its further development.

WP 2 **Cross-organizational use of financial resources**

This WP deals with the organizational perspective towards consolidating library budgets, third-party funds, and other financial resources to improve the cost monitoring of the OA transformation at research-performing organizations. WP 2 aims to develop a handout for scientific institutions on how to implement integrated information budgets; this will be prepared by an expert workshop, and accompanied by a close exchange with the community.

WP 3 **Further development and internationalization of a set of recommendations**

WP 3 continues activities of the Alliance of Science Organizations, who formulated recommendations for transformative agreements with publishers. These recommendations were published in November 2022 [124]. The objective of WP 3 is to further develop, spread, and internationalize this set of recommendations.

WP 4 **Process optimization**

This WP is dedicated to optimizing workflows related to handling publications and the associated metadata and invoices. WP 4 aims to improve the cooperation between institutions and publishers, and by focusing on electronic invoice processing and metadata management.

WP 5 **Transparency Initiative**

This WP addresses transparency in the OA transformation. Currently, the influence of confidential traditional subscription contracts is still at work. There is a need for action in this field to achieve more openness and transparency. The goal in WP 5 is to develop a concept for a national transparency initiative to promote sustainability in scientific publishing and to implement this approach as a standard.

WP6 **Competencies**

This WP deals with the organizational aspects of the OA transformation in libraries. WP 6 aims to develop flexible organizational structures and competence profiles for library professionals. In this, WP 6 especially focuses on the changes regarding the framework conditions as a result from the OA transformation.

WP 7 **Project management and public relations**

The WP ensures the project's management and organizes the overall internal as well as external communication work of the project.

8.4 Outlook

As of December 2022, the project is still in its early stages. Central to the project is the cooperation with other stakeholders on the national and international level. In the area of cost management, there will be a particularly close cooperation with the openCost project; Transform2Open was already featured in the openCost workshop in October 2022. Further, the Central Library of Forschungszentrum Jülich will contribute its experience of the Open Access Monitor Germany (OAM) for Germany [125]. Funded by the German Federal Ministry of Education and Research (BMBF), the OAM provides RPOs and funders with a freely available tool to analyze data on publications and related costs from German academic institutions. Additionally, the results of Transform2Open will further contribute to

the OA activities of the Alliance of Science Organizations in Germany. A central activity of the Alliance of Science Organizations is the internationally renowned DEAL. The Transform2Open project partners support DEAL and two of the members of the Transform2Open team are also involved in the DEAL working group. Furthermore, the Transform2Open results will also be presented and discussed in the Digital Information priority initiative of the Alliance of Science Organizations in Germany [126]. With this initiative, the science organizations in Germany cooperate on the challenges of digital transformation in science. The promotion of OA is a core activity of this initiative. The results of the Transform2Open project will be openly and transparently communicated, e. g. in workshops, webinars, and conference talks. All publications of the project will be published in OA, starting with the project proposal [49]. Transform2Open has thus embarked on its journey to contribute to a successful OA transformation in Germany, embedded in the national and international landscape of scientific publishing.

References

[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 (cit. on pp. 21, 46, 95, 98, 121).

[69] Bianca Schweighofer. "Projekt openCost auf dem 8. Bibliothekskongress 2022". de. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* (2022). Artwork Size: 1-10 Seiten Publisher: o-bib. Das offene Bibliotheksjournal / Herausgeber VDB, pp. 1–10. DOI: 10.5282/O-BIB/5862 (cit. on pp. 37, 96).

[70] German Science and Humanities Council. *Recommendations on the Transformation of Academic Publishing. Towards Open Access*. Position Paper Drs. 9477-22_en. Jan. 2022. DOI: 10.57674/0gtq-b603 (cit. on pp. 46, 48, 95).

[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 (cit. on pp. 47, 75, 94, 95, 121).

[122] Bernhard Mittermaier. “The German DEAL”. In: Open-Access-Tage 2020, Bielefeld (Germany), 15 Sep 2020 - 17 Sep 2020. Sept. 15, 2020. DOI: http://hdl.handle.net/2128/25702 (cit. on p. 96).

[123] Nina Leonie Weisweiler. “Im Spannungsfeld zwischen Zweckrationalität und Idealismus – Eine Analyse des Fachdiskurses zu Projekt DEAL mit Fokus auf den Begriff „Open Access“”. en. In: *Bibliothek Forschung und Praxis* 45.1 (Apr. 2021), pp. 163–183. DOI: 10.1515/bfp-2020-0116 (cit. on p. 96).

[124] Heinz Pampel et al. *Recommendations for Transformative Journal Agreements with Providers of Publishing Services. Guidelines of the priority initiative “Digital Information” of the Alliance of Science Organisations in Germany, against the background of the implementation of the Open Access Strategy 2021 – 2025 of the Alliance of Science Organisations in Germany*. en. Priority Initiative “Digital Information”, Alliance of Science Organizations in Germany, 2022. 9 pp. DOI: 10.48440/ALLIANZOA.046 (cit. on p. 96).

[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 (cit. on pp. 97, 113, 115).

[126] Steering Committee. *Shaping digital transformation in science. Mission statement 2018 - 2022*. en. Priority Initiative “Digital Information”, Alliance of Science Organizations in Germany, 2017. 17 pp. DOI: 10.2312/ALLIANZOA.016 (cit. on p. 98).

9

Spreading Publication Cost Information with the Electronic Journals Library (EZB)

Colin Sippl[1]
Universitätsbibliothek Regensburg

Silke Weisheit[2]
Universitätsbibliothek Regensburg

DOI: 10.3204/DESY-PROC-2023-01/09

Abstract

The Electronic Journals Library (EZB) presents e-journals to library users in a clearly structured, user-friendly interface and provides its member libraries with an efficient administrative tool for e-journal licences. Furthermore, the EZB offers information on more than 110,000 e-journals from all subject areas, including 75,000 freely available titles. Over 650 libraries and research institutions – mostly from German-speaking countries – jointly maintain the EZB data, ensuring that journal information is of a high standard and constantly updated.

This makes the EZB a reliable hub for high-quality, up-to-date data in the library world. It provides data for various user services delivering literature and research information. Within the scope of the openCost project, the EZB will be further extended by additional data on publication

[1]https://orcid.org/0000-0002-8503-1740
[2]https://orcid.org/0000-0002-2609-4274

costs and new functionalities for displaying and delivering these cost data via various existing EZB interfaces. These innovations include enhanced user information on whether their institution covers these publication costs, either via transformative agreements or via its publication funds. This is also being achieved by means of new journal categories, an OpenAPC integration, and better integration of transformative agreements via ESAC-IDs into EZB data. Following this path, the EZB is making a significant contribution to greater publication cost transparency.

9.1 Introduction

The EZB is a service by the University Library of Regensburg which provides information on scholarly e-journals [47]. Established in 1997, it currently offers information on more than 110,000 journals in all subject areas. In addition, 134,000 journals provided by aggregators are listed. The EZB also contains 75,000 journals which are accessible free of charge to everyone. Furthermore, participating libraries provide their users access to the journals they have subscribed to [127].

9.2 EZB User Interface

Via the institution-specific EZB front-end, users are given uniform access to e-journals and their full texts. The availability of full-text access is indicated by traffic-light symbols according to the license state at each member library as shown in Table 9.1.

	Free access to full-text articles.
	This institution has a license for this journal. Access to full-text articles is granted to users of this member library on the basis of a subscription.
	Only some of the published volumes are accessible as full-texts for the member library.
	No access to full-texts as the member library does not have a current subscription to this journal. However, tables of contents and abstracts are generally available free of charge.
€	For journals with an additional Euro symbol, individual articles can be accessed on a pay-per-view basis.

Table 9.1: EZB Traffic Light System

Beyond that, the EZB front-end provides users with journal metadata and information on publisher OA policies [128], including information offered by the Sherpa Romeo service [129] (Figure 9.1) Furthermore, various tools, such as the institution-specific availability check via the EZB Linking Service [130], are also being provided.

9.3 Cooperative Management of E-journals

More than 650 institutions, mostly from Germany, use the EZB to manage access and license information as well as metadata on e-journals. All the information on journals is maintained cooperatively, resulting in a high quality and topicality of the EZB data. EZB data is managed via a centralised interface: the EZB administration. It provides a convenient administration interface for the licenses of all participating institutions. In addition, journal packages used by multiple license participants are maintained centrally by negotiating institutions and transferred to all participating EZB user libraries via automated procedures. These journal packages are identified in the EZB via the unique EZB package ID (e.g. WIDEA for Wiley DEAL) and may contain sub-packages, so-called EZB collections, with another unique ID type (e.g. EZB-WIDEA-01707). A distinction is made between consortium packages, alliance licenses, national licenses, national consortia, aggregator packages and publisher packages [131]. Thus, all journals of a journal package can be retrieved and managed with the help of its unique ID. The EZB is interlinked with various digital services of other libraries and information institutions such as JOIN² [19], oa.finder [132, 133], or LIVIVO [134]. It provides specialised services as an important component of the infrastructure for the supply of scientific literature and information.

9.4 Role of the EZB in openCost

The aim of the openCost project is the creation of a technical infrastructure that enables free and easy access to publication cost data via standardised interfaces and metadata formats. Ultimately, this should lead to cost transparency at institutional, national and international level. With this goal in mind, the EZB will be expanded to include new functions for displaying publication costs, which will also be made available for further usage. For this purpose, an interface between OpenAPC [15] and the EZB will be implemented (Figure 9.2). In addition, institutions will be

Nature

Online Availability: Fulltexts available

To the fulltexts: Vol. 385, Iss. 6611 (1997) -

To the fulltexts: Nationallizenz : Vol. 1 (1869) - Vol. 462, Iss. 7276 (2009) (funded by the DFG)

(funded by the DFG)

Provided by:	University Library of Regensburg The full texts are free for the staff and students of the University of Regensburg (from within the Campus network). For access restrictions please see the Readme. Terms of use

Print Availability: Print version available too

print version(s) available

General information on the online edition:

Title:	Nature
Latest articles:	Newsfeed
Publisher / Provider:	Springer Nature Search publishers open access policy in SHERPA/RoMEO
Subject(s):	Biology, Biotechnology | Chemistry and Pharmacology | General Science | Medicine | Physics
Keyword(s):	Allgemeine Naturwissenschaften
Language(s):	English
E-ISSN(s):	1476-4687
P-ISSN(s):	0028-0836
ZDB-Number:	1413423-8
Fulltext online since:	Jg. 1, H. 1 (1869)
Homepage(s):	https://www.nature.com/
Type of Appearance:	Fulltext, online only
Pricetype:	subject to fee

Figure 9.1: Detailed View of a Journal Entry. (Illustration created by the authors, CC BY-NC-SA)

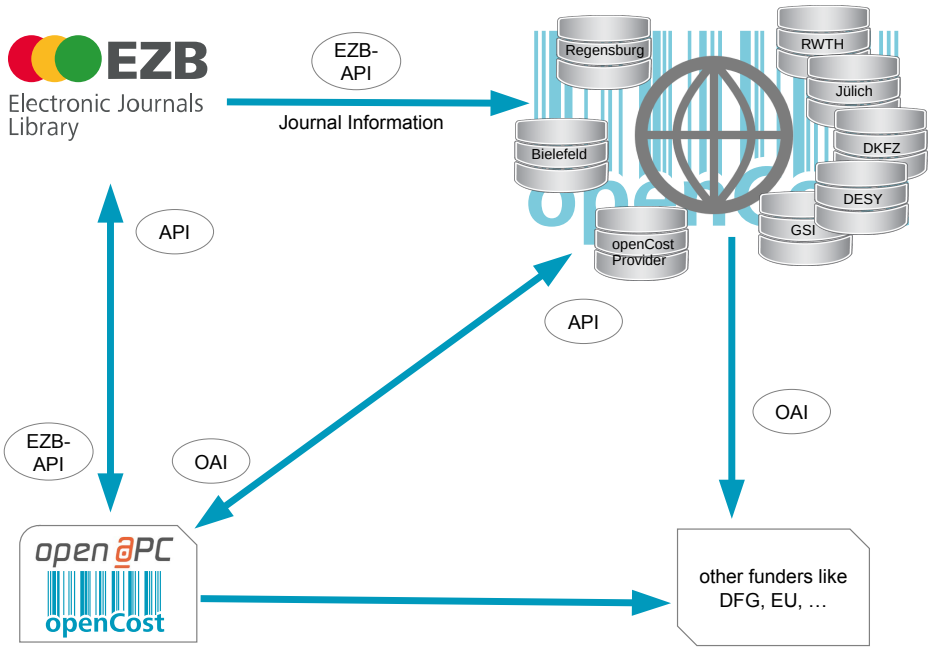

Figure 9.2: Role of the EZB in openCost. (Illustration by Alexander Wagner, CC BY-NC-SA)

provided with an option for entering information on the assumed publication costs for selected journals as well as institution-specific agreements, such as memberships. This information will be displayed in the EZB for the individual journals on an institution-specific basis and made available for further use (Figure 9.2). The EZB can thus be used as a central information platform for communicating OA information to researchers and external data services.

9.5 First EZB Extensions in openCost

The recent EZB upgrades and renewals include improved data links to external services, such as OpenAPC and DOAJ [9], as well as enhanced classifications of journals among the various EZB interfaces. Furthermore, better support for transformative agreements will improve publication cost transparency and interoperability of the EZB services. The new features are described in more detail below.

9.5.1 OpenAPC Connection

Providing publication cost data to users has become an important task for libraries. Eventually, users will need information not only on absolute costs, but also on available funding. The former involve datasets on fees paid for open access publishing. Therefore, data exchange between the OpenAPC service and the EZB is being established. To facilitate this data exchange, member libraries are required to provide their ROR-IDs [135]. For this purpose, an input field will be created in the EZB administration where each member library can enter its ROR-ID. The OpenAPC publication cost data will then be queried via the ROR-ID and subsequently integrated into the EZB administration interface and, at a later stage, into the user front-end. Thereby, the average APCs are displayed on an institution-specific and global level. With the help of this information, users can realistically estimate APC costs.

9.5.2 Journal Categories

The EZB journals will get new labels with special categories. Each journal category is a unique keyword that may refer to either an OA type or to an index the journal in question is listed in. These journal categories, which will appear in the EZB user front-end as keyword-like labels, can be used as additional search criteria and also help improve the EZB data export. Moreover, every journal category is provided with a definition, thereby creating a comprehensive glossary of all journal categories available in the EZB. With the help of journal categories, entire lists of journals can be displayed in the user interface or exported as KBART files. In addition, the journal categories serve as a technical basis for mapping journal funding conditions in the future. Furthermore, it is planned to display the OA funding conditions for journals in the EZB on an institution-specific basis. This will initially be implemented as an example for users of the institutions participating in the openCost project.

The sources for these categories include lists and databases that contain information on publication costs. These data sources need to be reliable (i.e., regular updates, transparent data collection and approval by the openCost and EZB community). Each category is defined by its name, a definition, and its source. So far, four provisional journal categories have been defined Table 9.2.

Journal category	Definition (preliminary)	Auto	Source
Indexed in DOAJ	The Directory of Open Access Journals (DOAJ) indexes and provides access to high quality, open access, peer-reviewed journals. All DOAJ journals are automatically tagged in the EZB.	✓	https://doaj.org/
Journal without APC	Full OA journal without publication fees.	✓	https://doaj.org/
Mirror Journal	A mirror journal is a fully open access version of an existing subscription journal, with the same editorial board, aims and scope, peer review processes and policies and an editorial board with at least 50% of the same members. The journal may have a similar name as the subscription title, but it must have a different ISSN.	✓	https://doi.org/10.26165/JUELICH-DATA/JRBK07
Subscribe to Open (S2O)	S2O allows publishers to convert journals from subscriptions to OA, one year at a time. Using S2O, a publisher offers a journal's current subscribers continued access. If all current subscribers participate in the S2O offer (simply by not opting out) the publisher opens the content covered by that year's subscription.	no	https://subscribetoopencommunity.org/

Table 9.2: Current EZB Journal Categories

Constantly updated data sets like DOAJ allow for automated journal category assignments. Thus, there are two types of journal categories: ones that can be assigned automatically and ones that need to be assigned manually. Currently, the Subscribe to Open (S2O) category would be an instance of the latter, as there are no larger datasets with S2O journals available. For journal categories requiring manual assignment, a selection option will be created in the EZB administration.

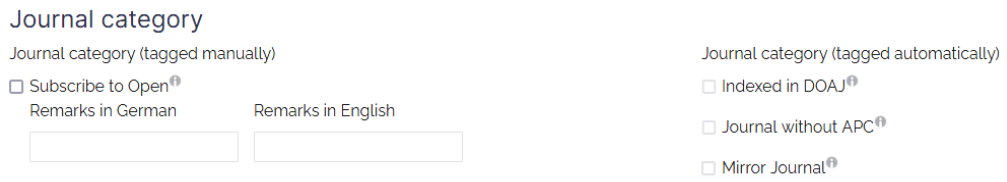

Figure 9.3: Journal Category Management in the EZB Administration. (Illustration created by the authors, CC BY-NC-SA)

If additional information on a journal category assignment needs to be included in an internal remark, it can be entered in an input field. In addition to the journal categories mentioned above, other journal categories are currently being considered, e.g., a category indicating whether a journal has flipped (either from OA to closed access or vice versa), whether it is a Diamond OA Journal, etc. New categories are jointly agreed upon by the openCost community.

9.5.3 Transformative Agreements

Within the scope of the openCost project, EZB collections will be labelled if they are part of transformative agreements. To identify the EZB packages and collections that map transformative agreements, the EZB administration will be extended by a check mark option as well as an input option to include the ESAC-ID [136], if available. This allows package managers to label their EZB collections accordingly. In addition, the terms and conditions of the transformative agreements will be shown in the user front-end so that users have immediate insight into the costs incurred. Thus, the EZB can provide its users with information on transformative agreements maintained in the EZB. This information will also be provided via the existing EZB interfaces.

9.6 Conclusion

The openCost enhancements to the EZB integrate information on publication costs and funding conditions into the EZB database and the user interface for the first time. Through the wide reach of the EZB and its various interfaces and data delivery services, this information is passed on to the community and can be used immediately. With these changes, the EZB is already making a major contribution to further cost transparency. In addition, the newly created journal

categories clearly define terms related to publishing and help to further increase the quality of the data provided. This will save a considerable amount of time and effort for both researchers and institutions, as the funding conditions don't have to be checked manually in each case. Furthermore, this would allow a quick comparison of the costs of individual institutions at journal level and improve the negotiating positions of libraries and consortia.

References

[9] *Directory of Open Access Journals. DOAJ*. en. URL: https://doaj.org/ (visited on 2022-12-06) (cit. on pp. 3, 105, 114).

[15] Bielefeld University Library and DINI working group Electronic Publishing. *OpenAPC*. en. URL: https://openapc.net/ (visited on 2022-12-06) (cit. on pp. 8, 16, 103).

[19] JOIN². *JOIN² - Just anOther INvenio INstance To Join*. en. URL: https://join2.de/ (visited on 2022-12-08) (cit. on pp. 11, 12, 38, 103).

[47] Universitätsbibliothek Regensburg. *Elektronische Zeitschriftenbibliothek (EZB)*. de. URL: https://ezb.ur.de/ (visited on 2022-12-06) (cit. on pp. 17, 102).

[127] Electronic Journals Library. *About the EZB*. en. URL: https://ezb.ur.de/about.phtml?lang=en (visited on 2022-12-07) (cit. on p. 102).

[128] Electronic Journals Library. *EZB Tool to find Open Access Rights*. en. URL: http://ezb.ur.de/services/oa.phtml?lang=en (visited on 2022-12-06) (cit. on p. 103).

[129] JISC. *Welcome to Sherpa Romeo - v2.sherpa*. en. URL: https://v2.sherpa.ac.uk/romeo/ (visited on 2022-12-06) (cit. on p. 103).

[130] Electronic Journals Library. *EZB Linking Service*. en. Dec. 2022. URL: http://ezb.ur.de/services/linkingdienst.phtml?lang=en (cit. on p. 103).

[131] Deutsche Forschungsgemeinschaft. *Über nationale Lizenzen — Nationallizenzen*. de. URL: https://www.nationallizenzen.de/ueber-nationallizenzen (visited on 2022-12-09) (cit. on p. 103).

[132] Bielefeld University Library and oa.network. *oa.finder*. en, de. URL: https://finder.open-access.network/ (visited on 2022-12-08) (cit. on p. 103).

[133] Nina Schönfelder. *Der ‚oa.finder' des Projekts ‚open-access.network'*. Presentation. 17 pp. DOI: 10.11576/kwi-977 (cit. on p. 103).

[134] ZB Med. *LIVIVO - The Search Portal for Life Sciences*. en. URL: https://www.livivo.de/ (visited on 2022-12-08) (cit. on p. 103).

[135] *Research Organization Registry (ROR)*. URL: https://ror.org/ (visited on 2022-12-08) (cit. on pp. 106, 114).

[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. 108, 115).

10

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

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

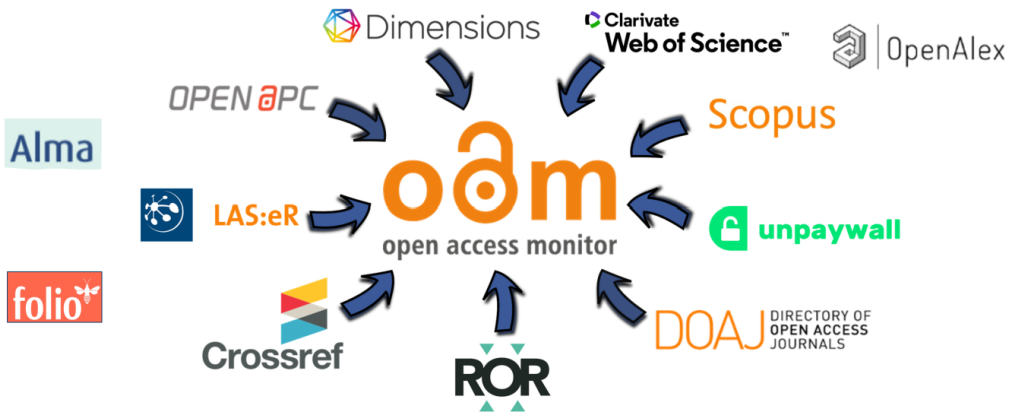

Figure 10.1: The Open Access Monitor Source Systems

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.

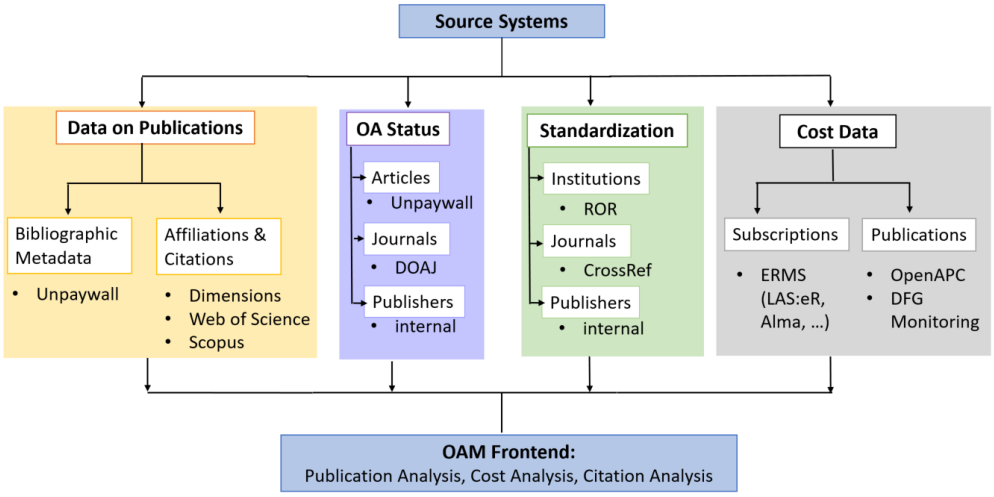

Figure 10.2: The OAM Database

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.

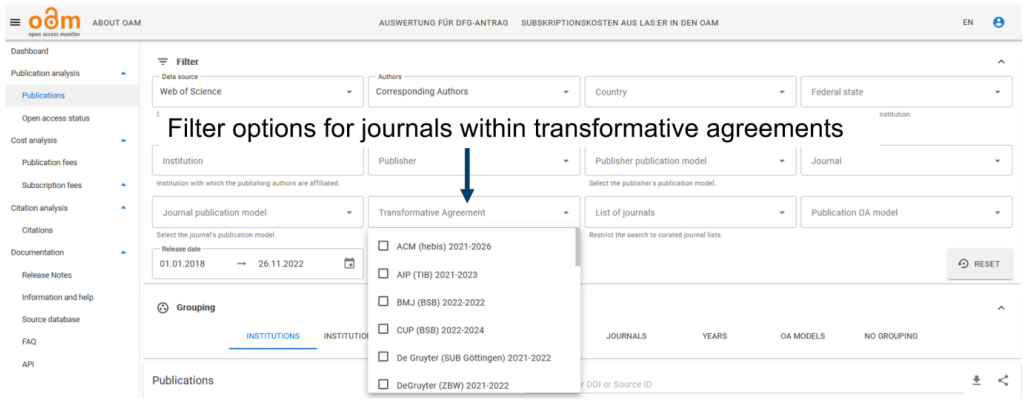

Figure 10.3: Filter options for journals within transformative agreements in the OAM user interface

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.

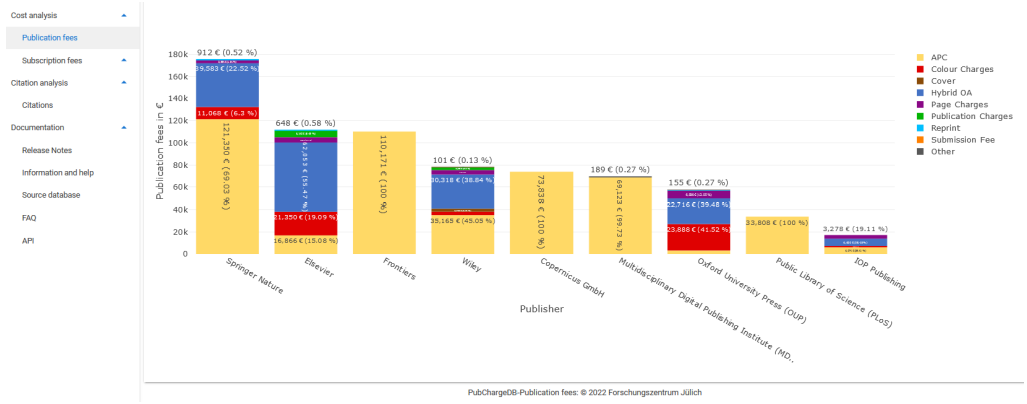

Figure 10.4: Institutional publication costs broken down by cost type and publisher in the OAM user interface

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,

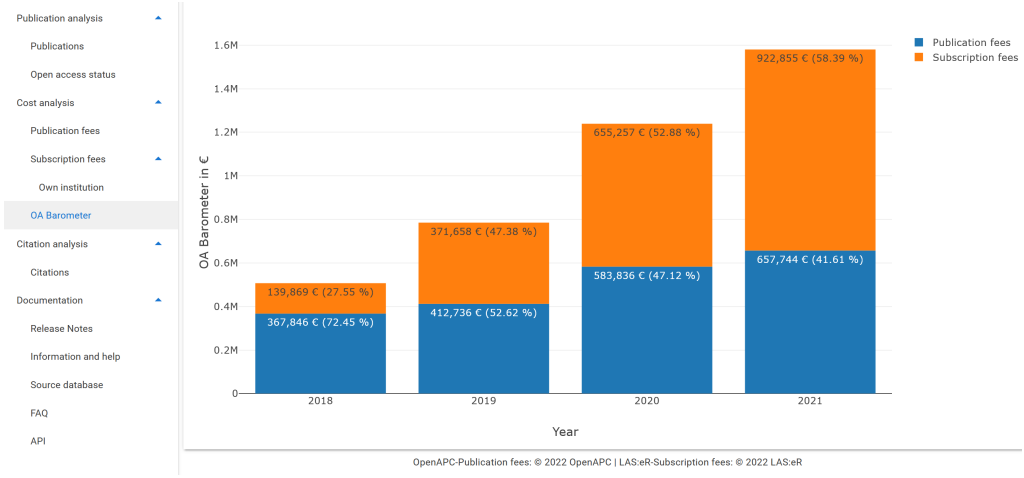

Figure 10.5: Institutional subscription fees vs. publication fees broken down by year.

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

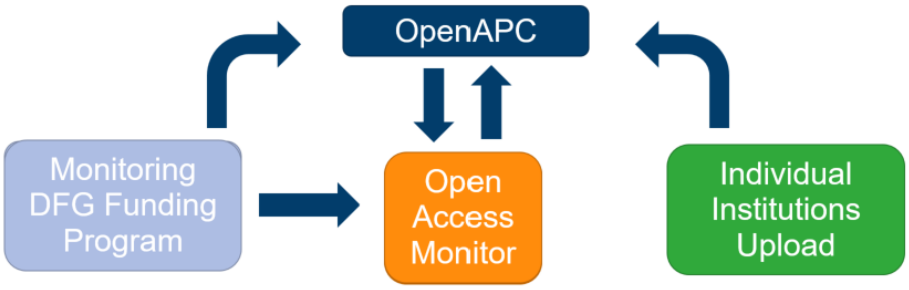

Figure 10.6: Data sharing cycle for total publication cost transparency

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, 105, 114).

[37] openCost. *openCost: Projekt*. URL: https://www.opencost.de/projekt/ (visited on 2022-11-27) (cit. on pp. 14, 16, 112).

[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, 112).

[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 (cit. on pp. 21, 46, 95, 98, 121).

[50] Deutsche Forschungsgemeinschaft. *Open Access Publication Funding: Guidelines and Supplementary Instructions*. DFG form 12.21. 2022 (cit. on pp. 23, 112, 115).

[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, 112).

[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 (cit. on pp. 47, 75, 94, 95, 121).

[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, 120).

[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 (cit. on pp. 97, 113, 115).

[135] *Research Organization Registry (ROR)*. URL: https://ror.org/ (visited on 2022-12-08) (cit. on pp. 106, 114).

[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. 108, 115).

[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. 112).

[138] Deutsche Forschungsgemeinschaft. *DFG. Deutsche Forschungsgemeinschaft*. URL: https://www.dfg.de/en/index.jsp (visited on 2022-11-27) (cit. on p. 112).

[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 (cit. on p. 113).

[140] OurResearch. *Unpaywall: Data feed*. URL: https://unpaywall.org/products/data-feed (visited on 2022-11-27) (cit. on p. 113).

[141] Digital Science & Research Solutions Ltd. *Dimensions*. URL: https://www.dimensions.ai/ (visited on 2022-11-27) (cit. on p. 113).

[142] Clarivate. *Web of Science*. URL: https://www.webofscience.com/wos/woscc/basic-search (visited on 2022-11-27) (cit. on p. 113).

[143] Elsevier B.V. *Scopus*. URL: https://www.scopus.com (visited on 2022-11-27) (cit. on p. 113).

[144] Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz). *LAS:eR*. URL: https://laser.hbz-nrw.de/ (visited on 2022-11-27) (cit. on p. 113).

[145] ExLibris. *Alma*. URL: https://www.exlibrisgroup.com/products/alma-library-services-platform/ (visited on 2022-11-27) (cit. on p. 113).

[146] The Open Library Foundation. *FOLIO*. URL: https://www.folio.org/ (visited on 2022-11-27) (cit. on p. 113).

[147] OurResearch. *OpenAlex*. URL: https://openalex.org/ (visited on 2022-11-27) (cit. on p. 113).

[148] *Crossref*. URL: https://www.crossref.org/ (visited on 2022-11-27) (cit. on p. 114).

[149] Digital Science & Research Solutions Ltd. *GRID*. URL: https://grid.ac/ (visited on 2022-11-27) (cit. on p. 114).

List of Figures

List of Tables

Bibliography

[1] Forschungszentrum Jülich. *Monitoring DFG-Förderprogramm „Open Access Publikationskosten“. Erläuterungen zum Datenschema*. Apr. 6, 2023. URL: https://www.fz-juelich.de/de/zb/downloads/open-science/monitoring-oa-publikationskosten-datenschema-erlaeuterungen-2023-01-25.pdf/@@download/file (cit. on pp. 2, 16, 23).

[2] Maxi Kindling et al. “Open-Access-Landesstrategien und der Weg zur offenen Wissenschaft – Möglichkeiten und Herausforderungen am Beispiel Berlin”. de. In: (2021). DOI: 10.5281/ZENODO.5535754 (cit. on p. 2).

[3] DEAL Konsortium. *About DEAL*. Apr. 6, 2023. URL: https://deal-konsortium.de/en/about-deal (cit. on p. 3).

[4] Open Library of Humanities. *Open Library of Humanities*. Apr. 6, 2023. URL: https://www.openlibhums.org/ (cit. on p. 3).

[5] SciPost Foundation. *SciPost*. Apr. 6, 2023. URL: https://scipost.org/ (cit. on p. 3).

[6] CERN. *Sponsoring Consortium for Open Access Publishing in Particle Physics. SCOAP³*. Apr. 6, 2023. URL: https://scoap3.org (cit. on pp. 3, 16).

[7] Public Library of Science. *PLOS Community Action Publishing*. Apr. 6, 2023. URL: https://plos.org/resources/community-action-publishing/ (cit. on p. 3).

[8] TIB Leibnitz Centre for Science and Technology University Library. *Konsortiale Open-Access-Lösungen aufbauen. KOALA*. Apr. 6, 2023. URL: https://projects.tib.eu/koala/en/ (cit. on p. 3).

[9] *Directory of Open Access Journals. DOAJ*. en. URL: https://doaj.org/ (visited on 2022-12-06) (cit. on pp. 3, 105, 114).

[10] Bernhard Mittermaier. "Informationsbudget". In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 9.4 (2022), pp. 1–17. DOI: 10.5282/o-bib/5864 (cit. on pp. 4, 46).

[11] German Science and Humanities Council. "Empfehlungen zur Transformation des wissenschaftlichen Publizierens zu Open Access". de. In: (2022). DOI: 10.57674/FYRC-VB61 (cit. on p. 4).

[12] Heinz Pampel. *Auf dem Weg zum Informationsbudget. Zur Notwendigkeit von Monitoringverfahren für wissenschaftliche Publikationen und deren Kosten*. de. Arbeitspapier. 2019. DOI: 10.2312/OS.HELMHOLTZ.006 (cit. on pp. 6, 17, 47, 72, 75).

[13] DINI AG Elektronisches Publizieren. *Dokumentation der Rechtsgrundlage für Veröffentlichungen in Repositorienmetadaten. Positionspapier*. de. DINI Schriften 20. Humboldt-Universität zu Berlin, 2022. DOI: 10.18452/23191 (cit. on p. 6).

[14] Deutsche Forschungsgemeinschaft. *Funding Programme: "Open Access Publication Funding"*. URL: https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/open_access_publikationskosten/ (visited on 2022-12-13) (cit. on pp. 8, 16, 21, 22, 47, 49).

[15] Bielefeld University Library and DINI working group Electronic Publishing. *OpenAPC*. en. URL: https://openapc.net/ (visited on 2022-12-06) (cit. on pp. 8, 16, 103).

[16] Open Archives Initiative. *Protocol for Metadata Harvesting. Protocol Version 2.0 of 2002-06-14*. Jan. 8, 2015. URL: http://www.openarchives.org/OAI/openarchivesprotocol.html (visited on 2023-04-06) (cit. on pp. 8, 14, 15, 19).

[17] Dublin Core™ Metadata Initiative. *XML schemas to support the Guidelines for implementing Dublin Core™ in XML*. Apr. 6, 2023. URL: https://www.dublincore.org/schemas/xmls/ (cit. on p. 8).

[18] DataCite. *DataCite Metadata Schema*. Apr. 6, 2023. URL: https://schema.datacite.org/ (cit. on p. 8).

[19] JOIN². *JOIN² - Just anOther INvenio INstance To Join*. en. URL: https://join2.de/ (visited on 2022-12-08) (cit. on pp. 11, 12, 38, 103).

[20] A. Wagner. "Veröffentlichungsdatenbank und Volltextrepositorium". In: *BIT online* 14 (2011), pp. 45–48 (cit. on p. 12).

[21] A. Wagner. "Ein neues JUWEL?: Publikationsmanagement für Wissenschaft und Administration". de. In: 33. Arbeits- und Fortbildungstagung der Arbeitsgemeinschaft der Spezialbibliotheken e.V. - Sektion 5 im Deutschen Bibliotheksverband, Jülich. 2011 (cit. on p. 12).

[22] JOIN². *JOIN² Instanzen*. 2022. URL: https://join2.de/Main/Instanzen (visited on 2022-12-06) (cit. on pp. 12, 22, 38).

[23] Alexander Wagner and Robert Thiele. "Invenio@HGF – Collaborative repository infrastructure". In: Open Repositories 2014, Helsinki (Finland), 8 Jun 2014 - 13 Jun 2014. June 8, 2014. DOI: 10.3204/DESY-2014-02793 (cit. on p. 12).

[24] Alexander Wagner. "JOIN² - What it is and what it does". In: 2016. DOI: 10.3204/PUBDB-2016-01331 (cit. on p. 12).

[25] Alexander Wagner. "Publication database". In: EUCALL Workshop: User Access at Advanced Laser Light Sources, Dolní Břežany (Czechia), 21 Sep 2017 - 22 Sep 2017. Sept. 21, 2017. DOI: 10.3204/PUBDB-2017-10906 (cit. on p. 12).

[26] Laura Baracci, Alexander Wagner, and JOIN2. "JOIN² : A Publication Database and Repository based on Invenio". In: XXVII International Symposium on Nuclear Electronics & Computing, Budva Becici (Montenegro), 30 Sep 2019 - 4 Oct 2019. Sept. 30, 2019, pp. 51–57. DOI: 10.3204/PUBDB-2019-04497 (cit. on p. 12).

[27] Irina Filozova et al. "JINR Open Access Repository based on the JOIN² Platform". In: *Supplementary Proceedings of the XXII International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020)*. Vol. 2790. Data Analytics and Management in Data Intensive Domains 2020, Voronezh (Russia), 13 Oct 2020 - 16 Oct 2020. Aachen, Germany: RWTH Aachen, Oct. 13, 2020, pp. 142–155. DOI: 10.3204/PUBDB-2021-00818 (cit. on p. 12).

[28] Alexander Wagner et al. “JOIN² empowering a bibliographic infrastructure to support scientists”. In: CERN Workshop on Innovations in Scholary Communication, Geneva (Switzerland), 17 Jun 2015 - 19 Jun 2015. June 17, 2015. DOI: 10.3204/PUBDB-2015-02188 (cit. on p. 12).

[29] Alexander Wagner. “JOIN² – A scientists toolbox”. In: Dubna (Russian Federation), 13 Jan 2018 - 21 Jan 2018. Jan. 13, 2018. DOI: 10.3204/PUBDB-2018-00618 (cit. on p. 12).

[30] Alexander Wagner. “JuSER - Autorenhandling”. In: DINI-/Helmholtz-Workshop „Autorenidentifikation am Beispiel von ORCID“, , 15 Oct 2012. Oct. 15, 2012. DOI: 10.3204/PUBDB-2017-00856 (cit. on pp. 12, 13).

[31] Alexander Wagner. “Authority Control in Invenio”. In: 2nd Invenio User Group Workshop, Jülich (Germany), 18 Nov 2013 - 20 Nov 2013. Nov. 18, 2013. DOI: 10.3204/PUBDB-2017-00870 (cit. on pp. 12, 13).

[32] Rebecca Mehl et al. *Repositorium, Publikationsdatenbank und Katalog: Integrierte Lösungen für kleine Bibliotheken*. Tech. rep. 2018, p. 6. DOI: 10.3204/PUBDB-2019-01442 (cit. on p. 12).

[33] National Library of Medicine, ed. *Pubmed Central*. URL: https://www.ncbi.nlm.nih.gov/pmc (cit. on p. 13).

[34] Cornell University, ed. *arXiv*. URL: https://arxiv.org (cit. on p. 13).

[35] CERN. *Invenio*. URL: https://invenio-software.org/about/ (cit. on p. 14).

[36] Alexander Wagner. “Invenio @ HGF - Technical background”. In: Invenio Developers Forum, Geneva (Switzerland), 16 Jun 2013. June 16, 2013. DOI: 10.3204/PUBDB-2017-00861 (cit. on p. 14).

[37] openCost. *openCost: Projekt*. URL: https://www.opencost.de/projekt/ (visited on 2022-11-27) (cit. on pp. 14, 16, 112).

[38] Alexander Wagner. “Publikationsfonds”. In: Open Access Tage 2011, Regensburg (Germany), 5 Oct 2011 - 5 Oct 2011. Oct. 5, 2011. DOI: 10.3204/PUBDB-2017-00849 (cit. on pp. 14, 17).

[39] Alexander Wagner. “Publizieren ist nicht genug”. In: vol. 22. Schriften des Forschungszentrums Jülich. Reihe Bibliothek / library. 7. Konferenz der Zentralbibliothek des Forschungszentrums Jülich, Jülich (Germany), 14 Jun 2016 - 16 Jun 2016. Jülich: Zentralbibliothek Forschungszentrum Jülich, June 14, 2016, pp. 23–47. DOI: 10.3204/PUBDB-2016-01819 (cit. on pp. 14, 15, 22).

[40] Alexander Wagner. “Article Processing Charges and OpenAPC”. In: 4th Invenio User Group Workshop, Garching (Germany), 21 Mar 2017 - 24 Mar 2017. Mar. 21, 2017. DOI: 10.3204/PUBDB-2017-01358 (cit. on pp. 14, 16, 17).

[41] Alexander Wagner et al. “APC-Verwaltung im institutionellen Repositorium”. de. In: 107. Deutscher Bibliothekartag - offen & vernetzt, Berlin (Germany), 12 Jun 2018 - 15 Jun 2018. June 12, 2018. DOI: 10.3204/PUBDB-2018-02211 (cit. on p. 14).

[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, 112).

[43] Gernot Deinzer et al. “Including Open Access Information in institutional repositories”. In: CERN Workshop on Innovations in Scholary Communication, Geneva (Switzerland), 21 Jun 2017 - 23 Jun 2017. June 21, 2017. DOI: 10.3204/PUBDB-2017-05961 (cit. on pp. 16, 37).

[44] Gernot Deinzer et al. “Including Open Access Information In Institutional Repositories”. In: Open Access Tage 2017, Dresden (Germany), 11 Sep 2017 - 13 Sep 2017. Sept. 11, 2017. DOI: 10.5281/zenodo.887318 (cit. on p. 16).

[45] openCost. *openCost*. URL: https://www.opencost.de (visited on 2022-12-13) (cit. on pp. 16, 37, 46).

[46] Dagmar Sitek, Alexander Wagner, and Corinna Mescherowsky. “Open-Access-Publikationskosten @ JOIN²”. In: Open-Access-Tage 2022, Bern (Switzerland), 19 Sep 2022 - 21 Sep 2022. Sept. 19, 2022 (cit. on p. 16).

[47] Universitätsbibliothek Regensburg. *Elektronische Zeitschriftenbibliothek (EZB)*. de. URL: https://ezb.ur.de/ (visited on 2022-12-06) (cit. on pp. 17, 102).

[48] Alexander Wagner. “APC-Verwaltung im institutionellen Repositorium”. de. In: *GMS Medizin, Bibliothek, Information* 18.3 (Sept. 10, 2018), Doc21. DOI: 10.3205/mbi000422 (cit. on p. 19).

[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 (cit. on pp. 21, 46, 95, 98, 121).

[50] Deutsche Forschungsgemeinschaft. *Open Access Publication Funding: Guidelines and Supplementary Instructions*. DFG form 12.21. 2022 (cit. on pp. 23, 112, 115).

[51] *Open Access Policy of the Helmholtz Association. Version 1.0. Approved in the 119th General Assembly of the Helmholtz Association on 20-21 September 2022*. 2016. 2 pp. DOI: 10.48440/os.helmholtz.018 (cit. on p. 24).

[52] *Helmholtz Open Science Policy. Version 1.0. Approved in the 119th General Assembly of the Helmholtz Association on 20-21 September 2022*. 2022. 9 pp. DOI: 10.48440/os.helmholtz.056 (cit. on p. 24).

[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, 112).

[54] The Open Knowledge Foundation. *Open Database License (ODbL) v1.0*. Nov. 30, 2022. URL: https://opendatacommons.org/licenses/odbl/1-0/ (visited on 2022-11-30) (cit. on p. 32).

[55] *INTACT project*. 2022. URL: https://gepris.dfg.de/gepris/projekt/274983645?language=en (visited on 2022-11-30) (cit. on p. 32).

[56] OpenAPC. *APC data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/apc_de.csv (visited on 2022-11-30) (cit. on p. 32).

[57] OpenAPC. *TA data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/transformative_agreements/transformative_agreements.csv (visited on 2022-11-30) (cit. on p. 32).

[58] OpenAPC. *BPC data set on GitHub*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/data/bpc.csv (visited on 2022-11-30) (cit. on p. 32).

[59] OpenAPC. *OpenAPC GitHub repository*. 2022. URL: https://github.com/OpenAPC/openapc-de (visited on 2022-11-30) (cit. on p. 33).

[60] OpenAPC. *OpenAPC first release*. 2022. URL: https://github.com/OpenAPC/openapc-de/tree/v1.0.0 (visited on 2022-11-30) (cit. on p. 33).

[61] OpenAPC. *OpenAPC README file*. 2022. URL: https://github.com/OpenAPC/openapc-de/blob/master/README.md (visited on 2022-11-30) (cit. on p. 33).

[62] OpenAPC. *OpenAPC data schemas*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/schema (visited on 2022-12-06) (cit. on p. 34).

[63] Dirk Pieper and Christoph Broschinski. "OpenAPC: a contribution to a transparent and reproducible monitoring of fee-based open access publishing across institutions and nations". In: *Insights the UKSG journal* 31 (2018), p. 39. DOI: 10.1629/uksg.439 (cit. on pp. 34, 47).

[64] OpenAPC. *OpenAPC cost definition*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/Data-Submission-Handout#definition-of-costs (visited on 2022-11-30) (cit. on p. 35).

[65] OpenAPC. *OpenAPC handout on DEAL data*. 2022. URL: https://github.com/OpenAPC/openapc-de/wiki/Handreichung-Dateneingabe-Transformationsvertr%C3%A4ge-%28DEAL-Wiley-und--Springer-Nature%29-ab-Berichtsjahr-2020 (visited on 2022-11-30) (cit. on p. 35).

[66] OpenAPC. *OpenAPC blog*. 2022. URL: https://openapc.github.io/ (visited on 2022-11-30) (cit. on p. 35).

[67] OpenAPC. *OpenAPC treemaps*. 2022. URL: https://treemaps.openapc.net/ (visited on 2022-11-30) (cit. on p. 36).

[68] OpenAPC. *OpenAPC OLAP server*. 2022. URL: https://olap.openapc.net/ (visited on 2022-11-30) (cit. on p. 36).

[69] Bianca Schweighofer. “Projekt openCost auf dem 8. Bibliothekskongress 2022”. de. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* (2022). Artwork Size: 1-10 Seiten Publisher: o-bib. Das offene Bibliotheksjournal / Herausgeber VDB, pp. 1–10. DOI: 10.5282/O-BIB/5862 (cit. on pp. 37, 96).

[70] German Science and Humanities Council. *Recommendations on the Transformation of Academic Publishing. Towards Open Access*. Position Paper Drs. 9477-22_en. Jan. 2022. DOI: 10.57674/0gtq-b603 (cit. on pp. 46, 48, 95).

[71] openCost. *openCost auf dem Bibliothekskongress: Ergebnisse des Workshops*. URL: https://www.opencost.de/allgemein/ergebnisse-des-workshops/ (visited on 2022-12-13) (cit. on p. 46).

[72] *Workshop “openCost: the road to publication cost transparency”*. URL: https://www.opencost.de/allgemein/opencost-workshop/ (visited on 2022-12-13) (cit. on p. 46).

[73] Anita Eppelin et al. “Umgang mit Open-Access-Publikationsgebühren – die Situation in Deutschland in 2010”. In: *GMS Medizin - Bibliothek - Information* 12.1-2 (2012). DOI: 10.3205/MBI000240 (cit. on p. 46).

[74] Ralf Schimmer. “Open Access und die Re-Kontextualisierung des Bibliothekserwerbungsetats”. In: *Bibliothek Forschung und Praxis* 36.3 (2012), pp. 293–299. DOI: 10.1515/bfp-2012-0038 (cit. on p. 46).

[75] Stuart Lawson. “’Total cost of ownership’ of scholarly communication. managing subscription and APC payments together”. In: *Learned Publishing* 28.1 (2015), pp. 9–13. DOI: 10.1087/20150103 (cit. on p. 47).

[76] Irene Barbers, Nadja Kalinna, and Bernhard Mittermaier. “Data-Driven Transition: Joint Reporting of Subscription Expenditure and Publication Costs”. In: *Publications* 6.2 (2018). DOI: 10.3390/publications6020019 (cit. on p. 47).

[77] Andrew Gray. “Considering Non-Open Access Publication Charges in the “Total Cost of Publication””. In: *Publications* 3.4 (2015), pp. 248–262. DOI: 10.3390/publications3040248 (cit. on p. 47).

[78] Stephen Pinfield, Jennifer Salter, and Peter A. Bath. "The "total cost of publication" in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions". In: *Journal of the Association for Information Science and Technology* 67.7 (2016), pp. 1751–1766. DOI: 10.1002/asi.23446 (cit. on pp. 47, 72, 75, 87).

[79] Bernhard Mittermaier and Irene Barbers. "Zeitschriftenmonitoring". In: *Bibliotheken. Wegweiser in die Zukunft. : Projekte und Beispiele*. Ed. by Erda Lapp et al. Berlin: Simon Verlag für Bibliothekswissen, 2021, pp. 346–359. DOI: http://hdl.handle.net/2128/22746 (cit. on pp. 47, 75).

[80] Stuart Lawson and Ben Meghreblian. "Journal subscription expenditure of UK higher education institutions". In: *F1000Research* 3 (2015), p. 274. DOI: 10.12688/f1000research.5706.3 (cit. on p. 47).

[81] Arbeitsgruppe Open Access in der Schwerpunktinitiative „Digitale Information" der Allianz der Deutschen Wissenschaftsorganisationen. *Open Access gestalten, Transparenz schaffen! Appell zur Offenlegung von Publikationsgebühren*. 2016. DOI: 10.2312/ALLIANZOA.013 (cit. on p. 47).

[82] Christoph Bruch et al. *Empfehlungen zur Open-Access-Transformation. Strategische und praktische Verankerung von Open Access in der Informationsversorgung wissenschaftlicher Einrichtungen*. Recommendations. Ad-hoc-AG Open-Access-Gold im Rahmen der Schwerpunktinitiative „Digitale Information" der Allianz der deutschen Wissenschaftsorganisationen, 2016. DOI: 10.3249/ALLIANZOA.011 (cit. on p. 47).

[83] Christian Gutknecht. "Transparenz von Subskriptionskosten in der Schweiz". In: *027.7 Zeitschrift für Bibliothekskultur* 4.1 (2016), pp. 26–32. DOI: 10.12685/027.7-4-1-103 (cit. on p. 47).

[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 (cit. on pp. 47, 75, 94, 95, 121).

[85] DEAL Operations. *Here's the deal: Reduce and restructure costs*. URL: https://web.archive.org/web/20221128222942/https://deal-operations.de/en/here-is-the-deal/restructure-costs (visited on 2022-12-13) (cit. on p. 47).

[86] Gali Halevi and Samantha Walsh. “Faculty Attitudes Towards Article Processing Charges for Open Access Articles”. In: *Publishing Research Quarterly* 37.3 (2021), pp. 384–398. DOI: 10.1007/s12109-021-09820-x (cit. on p. 51).

[87] Florian Aumeier and Ingrid Heinen. “Inhouse-Lösung für das Jülicher Electronic Resource Management System”. In: *Bibliotheksdienst* 41.3 (2007). DOI: 10.1515/bd.2007.41.3.322 (cit. on p. 54).

[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, 120).

[89] Antoine Blanchard, Diane Thierry, and Maurits van der Graaf. *Retrospective and prospective study of the evolution of APC costs and electronic subscriptions for French institutions*. Research rep. Comité pour la science ouverte, Dec. 2022. DOI: 10.52949/26 (cit. on p. 63).

[90] Norwegian Ministry of Education and Research. *Norwegian national guidelines for Open Access*. Policy. Aug. 2017. 2 pp. (cit. on p. 64).

[91] SIKT. *CRISTIN – National CRIS. Reporting instructions (NVI)*. Mar. 2019. URL: https://www.cristin.no/english/resources/reporting-instructions/ (cit. on p. 64).

[92] Nina Karlstrøm, Lars Wenaas, and Vidar Røeggen. *Overgang til åpen publisering i Norge. Status for åpen publisering og kartlegging av kostnader*. 32 pp. (cit. on p. 64).

[93] The Research Council of Norway. *Support for Open Access Publication for the Financial Year 2021. STIM-OA*. Apr. 2022. URL: https://www.forskningsradet.no/en/call-for-proposals/2022/support-open-access-publication-2021/ (cit. on p. 65).

[94] Christoph Broschinski et al. *OpenAPC/openapc-de: Couperin Additions*. Version v4.79.23-3-3. Zenodo, Mar. 2023. DOI: 10.5281/zenodo.7707499 (cit. on p. 67).

[95] AT2OA². *Austrian transition to open access 2*. URL: https://www.at2oa.at/at2oa2_home.html (visited on 2022-12-06) (cit. on p. 72).

[96] Christof Capellaro et al. *Open-Access-Publikationsfonds. Einrichtung und Förderbedingungen*. de. 2019. DOI: 10.5281/ZENODO.2653724 (cit. on p. 72).

[97] Jessica Monaghan, Mithu Lucraft, and Katie Allin. *'APCs in the Wild'. Could Increased Monitoring and Consolidation of Funding Accelerate the Transition to Open Access?* White paper. 2020. DOI: 10.6084/M9.FIGSHARE.11988123.V4 (cit. on p. 72).

[98] AT2OA. *Empfehlung zur Buchung von Open-Access-Publikationskosten*. de. 2020. DOI: 10.5281/ZENODO.3945952 (cit. on pp. 72, 75, 79).

[99] AT2OA². *Überarbeitete Leitlinien für Verlagsverträge für die OA-Transformation im Rahmen von AT2OA2*. URL: https://www.at2oa.at/docs/at2oa2/Ueberarbeitete%20Leitlinien%20fuer%20Verlagsvertraege%20fuer%20die%20OA-Transformation%20im%20Rahmen%20von%20AT2OA2.pdf (visited on 2022-12-06) (cit. on p. 73).

[100] AT2OA². *Scholarly communications in transition: predatory and other phenomena in academia*. URL: https://in-transition.at/ (visited on 2022-12-06) (cit. on p. 74).

[101] Digital Science & Research Solutions Ltd. *Altmetric*. URL: https://www.altmetric.com/ (visited on 2022-12-06) (cit. on p. 74).

[102] Heinz Pampel. *Open Access an wissenschaftlichen Einrichtungen in Deutschland. Ergebnisse einer Erhebung im Jahr 2018*. de. Bericht. 2019. DOI: 10.2312/OS.HELMHOLTZ.005 (cit. on p. 75).

[103] Lisa Lovén. "Monitoring open access publishing costs at Stockholm University". In: *Insights the UKSG journal* 32 (2019). DOI: 10.1629/uksg.451 (cit. on p. 75).

[104] Claudia Frick and Christian Kaier. "Publikationskosten für Zeitschriftenartikel abseits von Open-Access-Publikationsfonds – Lost in Transformation?" de. In: *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 7.2 (2020), pp. 1–15. DOI: 10.5282/O-BIB/5586 (cit. on p. 75).

[105] Andre Bruns and Niels Taubert. "Investigating the Blind Spot of a Monitoring System for Article Processing Charges". In: *Publications* 9.3 (Sept. 2021), p. 41. DOI: 10.3390/publications9030041 (cit. on p. 75).

[106] *Bibliotheksstatistik.de*. URL: https://www.bibliotheksstatistik.de/ (visited on 2022-12-06) (cit. on p. 76).

[107] Timo Vilén et al. “Kirjoittajamaksut ja niiden seuranta: havaintoja ja kehitysehdotuksia. FinELibin Open APC -projektin loppuraportti”. In: *Kansalliskirjaston raportteja ja selvityksiä* 28 (2021) (cit. on pp. 85, 86).

[108] National Library of Sweden. *Open APC Sweden*. URL: https://www.kb.se/samverkan-och-utveckling/oppen-tillgang-och-bibsamkonsortiet/open-apc-sweden.html (visited on 2023-01-30) (cit. on p. 86).

[109] OpenAPC. *OpenAPC data schemas*. URL: https://github.com/OpenAPC/openapc-de/wiki/schema#openapc-data-set (visited on 2023-01-30) (cit. on p. 86).

[110] OpenAPC. *OpenAPC Finland*. URL: https://treemaps.openapc.net/apcdata/openapc/#institution/country=FIN (visited on 2023-01-30) (cit. on pp. 86, 89).

[111] *VIRTA Publication Information Service*. URL: https://www.csc.fi/en/-/virta-publication-information-service (visited on 2023-01-30) (cit. on p. 86).

[112] Open Science Coordination in Finland and Federation of Finnish Learned Societies. *Open access to scholarly publications. National Policy and executive plan by the research community in Finland for 2020 – 2025 (1)*. 2nd edition. Vol. 3. Responsible Research Series. Nov. 2019. DOI: 10.23847/isbn.9789525995343 (cit. on p. 86).

[113] National Library of Sweden. *Utgifter för vetenskaplig publicering*. URL: https://www.kb.se/samverkan-och-utveckling/oppen-tillgang-och-bibsamkonsortiet/utgifter-for-vetenskaplig-publicering.html (visited on 2023-01-30) (cit. on pp. 87, 89).

[114] *FinELib-konsortion vuosi 2021*. 2021. URL: https://finelib.fi/wp/wp-content/uploads/2022/06/FinELibin_vuosikertomus_2021.pdf (visited on 2023-01-30) (cit. on p. 87).

[115] Ministry of Education and Culture of Finland and Open Science and Research Initiative Finland. *Academic publisher costs in Finland 2010-2016*. Ministry of Education and Culture of Finland and its Open Science and Research Initiative 2014-2017 (cit. on p. 87).

[116] J. Forssbæck and L. Oxelheim. *The Multi-Faceted Concept of Transparency*. IFN Working Paper 1013. Research Institute of Industrial Economics (IFN), Stockholm, 2014. DOI: 10.2139/ssrn.2405951 (cit. on p. 87).

[117] N.L. Maron et al. *The Costs of Publishing Monographs: Toward a Transparent Methodology*. Research rep. Ithaka S+R., Feb. 2016. DOI: 10.18665/sr.276785 (cit. on p. 88).

[118] Ministry of Education and Culture of Finland. *Kustannukset*. URL: https://www.tiedejatutkimus.fi/fi/science-innovation-policy/open-science-and-research-indicators/indicators_content_4 (visited on 2023-01-30) (cit. on pp. 89, 90).

[119] Secretariat for the National Open Science and Research Coordination Federation of Finnish Learned Societies. *Open Science Monitoring Begins!* May 2022. URL: https://avointiede.fi/en/news/open-science-monitoring-begins (visited on 2023-01-30) (cit. on p. 89).

[120] *Fairdata.fi*. URL: https://www.fairdata.fi/en/ (visited on 2023-01-30) (cit. on p. 89).

[121] Bart de Langhe and Stefano Puntoni. "Leading with decision-driven data analytics". en. In: *MIT Sloan Management Review* 62.3 (2021) (cit. on p. 90).

[122] Bernhard Mittermaier. "The German DEAL". In: Open-Access-Tage 2020, Bielefeld (Germany), 15 Sep 2020 - 17 Sep 2020. Sept. 15, 2020. DOI: http://hdl.handle.net/2128/25702 (cit. on p. 96).

[123] Nina Leonie Weisweiler. "Im Spannungsfeld zwischen Zweckrationalität und Idealismus – Eine Analyse des Fachdiskurses zu Projekt DEAL mit Fokus auf den Begriff „Open Access"". en. In: *Bibliothek Forschung und Praxis* 45.1 (Apr. 2021), pp. 163–183. DOI: 10.1515/bfp-2020-0116 (cit. on p. 96).

[124] Heinz Pampel et al. *Recommendations for Transformative Journal Agreements with Providers of Publishing Services. Guidelines of the priority initiative "Digital Information" of the Alliance of Science Organisations in Germany, against the background of the implementation of the Open Access Strategy 2021 – 2025 of the Alliance of Science Organisations in Germany*. en. Priority Initiative "Digital Information",

Alliance of Science Organizations in Germany, 2022. 9 pp. DOI: 10.48440/ALLIANZOA.046 (cit. on p. 96).

[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 (cit. on pp. 97, 113, 115).

[126] Steering Committee. *Shaping digital transformation in science. Mission statement 2018 - 2022*. en. Priority Initiative "Digital Information", Alliance of Science Organizations in Germany, 2017. 17 pp. DOI: 10.2312/ALLIANZOA.016 (cit. on p. 98).

[127] Electronic Journals Library. *About the EZB*. en. URL: https://ezb.ur.de/about.phtml?lang=en (visited on 2022-12-07) (cit. on p. 102).

[128] Electronic Journals Library. *EZB Tool to find Open Access Rights*. en. URL: http://ezb.ur.de/services/oa.phtml?lang=en (visited on 2022-12-06) (cit. on p. 103).

[129] JISC. *Welcome to Sherpa Romeo - v2.sherpa*. en. URL: https://v2.sherpa.ac.uk/romeo/ (visited on 2022-12-06) (cit. on p. 103).

[130] Electronic Journals Library. *EZB Linking Service*. en. Dec. 2022. URL: http://ezb.ur.de/services/linkingdienst.phtml?lang=en (cit. on p. 103).

[131] Deutsche Forschungsgemeinschaft. *Über nationale Lizenzen — Nationallizenzen*. de. URL: https://www.nationallizenzen.de/ueber-nationallizenzen (visited on 2022-12-09) (cit. on p. 103).

[132] Bielefeld University Library and oa.network. *oa.finder*. en, de. URL: https://finder.open-access.network/ (visited on 2022-12-08) (cit. on p. 103).

[133] Nina Schönfelder. *Der ‚oa.finder' des Projekts ‚open-access.network'*. Presentation. 17 pp. DOI: 10.11576/kwi-977 (cit. on p. 103).

[134] ZB Med. *LIVIVO - The Search Portal for Life Sciences*. en. URL: https://www.livivo.de/ (visited on 2022-12-08) (cit. on p. 103).

[135] *Research Organization Registry (ROR)*. URL: https://ror.org/ (visited on 2022-12-08) (cit. on pp. 106, 114).

[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. 108, 115).

[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. 112).

[138] Deutsche Forschungsgemeinschaft. *DFG. Deutsche Forschungsgemeinschaft*. URL: https://www.dfg.de/en/index.jsp (visited on 2022-11-27) (cit. on p. 112).

[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 (cit. on p. 113).

[140] OurResearch. *Unpaywall: Data feed*. URL: https://unpaywall.org/products/data-feed (visited on 2022-11-27) (cit. on p. 113).

[141] Digital Science & Research Solutions Ltd. *Dimensions*. URL: https://www.dimensions.ai/ (visited on 2022-11-27) (cit. on p. 113).

[142] Clarivate. *Web of Science*. URL: https://www.webofscience.com/wos/woscc/basic-search (visited on 2022-11-27) (cit. on p. 113).

[143] Elsevier B.V. *Scopus*. URL: https://www.scopus.com (visited on 2022-11-27) (cit. on p. 113).

[144] Hochschulbibliothekszentrum des Landes Nordrhein-Westfalen (hbz). *LAS:eR*. URL: https://laser.hbz-nrw.de/ (visited on 2022-11-27) (cit. on p. 113).

[145] ExLibris. *Alma*. URL: https://www.exlibrisgroup.com/products/alma-library-services-platform/ (visited on 2022-11-27) (cit. on p. 113).

[146] The Open Library Foundation. *FOLIO*. URL: https://www.folio.org/ (visited on 2022-11-27) (cit. on p. 113).

[147] OurResearch. *OpenAlex*. URL: https://openalex.org/ (visited on 2022-11-27) (cit. on p. 113).

[148] *Crossref*. URL: https://www.crossref.org/ (visited on 2022-11-27) (cit. on p. 114).

[149] Digital Science & Research Solutions Ltd. *GRID*. URL: https://grid.ac/ (visited on 2022-11-27) (cit. on p. 114).

Contributors

Alexander Wagner, 11

Bernhard Mittermaier, 45, 95

Christian Kaier, 73
Christoph Broschinski, 31
Colin Sippl, 103

Gernot Deinzer, 1

Heinz Pampel, 95

Irene Barbers, 95, 113

Jens H. Aasheim, 65
Julia Bartlewski, 31

Kerstin Grossmaier-Stieg, 73

Lea Maria Ferguson, 95
Lisa-Marie Stein, 11

Silke Weisheit, 103

Timo Vilén, 87
Tobias Höhnow, 95

About the Project

About openCost

The openCost project was launched by Bielefeld University Library, the Deutsches Elektronen-Synchrotron, DESY (Hamburg) and Regensburg University Library and is funded by the German Research Foundation (DFG) (grant number 457354095).

Within the scope of openCost, we are creating a technical infrastructure that allows us to comprehensively record all costs involved in scientific publishing and subsequently make them freely accessible by means of standardized interfaces. This facilitates cost transparency on an institutional, national, and international level.

For this purpose, we are developing a standardized metadata schema to record, retrieve, and map all publication costs of a scientific institution in a structured form. This includes not only OA publication charges (APCs) but also costs from transformation contracts, memberships, etc.

We propose using the established OAI-PMH interface for an automatic exchange. As first applications, we are using it as a model on the publication servers of the universities of Bielefeld and Regensburg as well as the partner institutions of JOIN². Via OAI-PMH, service providers (e. g. aggregators and research funders) can harvest their publication outputs directly from the institutions. Within the framework of openCost, this will be realized by the OpenAPC service based at Bielefeld University Library.

A third focus of the project is the extension of the Electronic Journals Library (EZB) with special functions for displaying publication costs. To achieve this, we are implementing an interface between OpenAPC and the EZB. In addition, we

want to provide a function that allows institutions to enter information on funding options for specific journals as well as institution-specific agreements, such as memberships, in the EZB Administration. Thus, all participating institutions can use the EZB as a central platform for communicating OA information to researchers.

About the openCost workshop and the present conference proceedings

In October 2022, the expert workshop openCost: The Road to Publication Cost Transparency took place at the Deutsches Elektronen-Synchrotron, DESY in Hamburg as part of the DFG project openCost. Speakers from eight different countries contributed their perspectives in the fields of publication costs and cost transparency.

In addition to presenting the initial results of openCost itself, the main goal of the workshop was to provide an opportunity for knowledge exchange between national and international experts in the field. They presented their perspectives and shared their experiences to ensure openCost is internationally adoptable.

The expert workshop summarized the desiderata of the individual participants and the institutions they represent. These results served as a starting point to enhance and fine-tune the internal proposal for a metadata schema jointly developed by the openCost core members DESY Library for the JOIN² collaboration and the university libraries of Bielefeld (OpenAPC and the OpenAccess activities there) and Regensburg (EZB and the OpenAccess projects there).

The present conference proceedings summarize the contributions of the international and national experts as well as the results of the workshop.