

# Disclosing funding sources for open access publication fees: the Open APC initiative

Najko Jahn<sup>1</sup> and Marco Tullney<sup>2</sup>

<sup>1</sup>Bielefeld University Library, Bielefeld University, Bielefeld, Germany

<sup>2</sup>Technische Informationsbibliothek (TIB) - German National Library of Science and Technology, Hannover, Germany

## ABSTRACT

Publication fees in open access publishing hold a prominent place on the agenda of researchers, policy-makers, and academic publishers. This paper contributes to the evolving empirical basis on open access funding. It describes the Open APC initiative, in which German universities and research organizations share their expenditures for publication fees. As method, the initiative uses existing open data tools to aggregate and disseminate institutional spending on open access publication fees. In total, 29 German research organizations self-reported funding of 6,279 open access journal articles, which amounted to 8,039,339 €. The average payment for each article was 1,280 €, and the median payment 1,209 €. Our data-set comprises only 53 articles in hybrid journals. With an indexing coverage of 99 %, the findings reveal that the DOI agency CrossRef provides both comprehensive bibliographic coverage of the funded open access journal literature and disambiguated names of journal titles and publishing houses. We show that authority control of these bibliographic information is particularly relevant for the comparative study of the economical effects of open access publishing.

**Keywords:** Open access, open access journal, scholarly publishing, publication fees, article processing charges, science policy

## INTRODUCTION

### General Background

The rise of open access journals matches the increasing relevance of publication fees in academic publishing (Davis and Walters, 2011; Laakso and Björk, 2012; Pinfield, 2015). To cover publication fees, also referred to as article-processing charges (APC), authors tend to make use of funding that grant agencies or academic institutions provide (Suber, 2012). Yet, how and to what extent these research support activities are effective in terms of the number of supported articles and associated costs remains under debate.

One reason why studying how much is being spent on publication fees is in most cases difficult is that payments for these charges are fragmented across the budgets of funding agencies, research institutions, and libraries, or, if support is limited, are taken from personal budgets. Another key problem is that funding for open access journals using publication fees lacks transparency because the parties involved - authors, universities, funders, publishers - neither release information on who pays for what nor the costs of publishing (Björk and Solomon, 2014), a situation similar to the lack of transparency regarding journal subscriptions (Lawson and Meghreblian, 2015). It also remains unclear which factors contribute to price formation.<sup>1</sup> While fixed prices for individual articles are common, agreements between publishers and institutions often provide discounts, and publishers sometimes waive publication fees for authors from low-income countries (Björk and Solomon, 2012; Lawson, 2015c). Other factors leading to variable pricing schemes include submission or page charges (Björk and Solomon, 2012). Hybrid journals substantially add to this complexity of open access funding (Kingsley, 2014). These journals, allowing

<sup>1</sup>These might include article processing, impact, rejection rates, management and investment, and profit margins. See Noorden (2013) for a general discussion and Gumpenberger et al. (2012) and Björk and Solomon (2015) for discussions of journal impact and quality.

articles to be published immediately as open access after a charge was paid, rely both on subscriptions and publication fees as revenue sources. Although the uptake of open access through hybrid journals was described as lower and more expensive compared to that of fully open access journals (Björk and Solomon, 2014; Solomon and Björk, 2012), this model has gained attention through recent science policies, notably because of open access policies from the UK (Pinfield, 2015).

Previous empirical studies asking how much was being spent for publishing in open access journals confirmed that authors make use of funding or institutional support for covering publication fees. In 2010, the SOAP project carried out a comprehensive survey in which 9,645 authors from various disciplines provided insights into how they financed publication fees (Dallmeier-Tiessen et al., 2011). The majority of the respondents used institutional support structures or research funding. By contrast, 12 % paid publication fees individually. These findings are consistent with that of other surveys, adding that sources of funding exists mostly in higher income countries, mainly to support research articles in the bio- and physical sciences (Solomon and Björk, 2011). Personal budgets, on the other hand, are likely used to cover low price publication fees (Björk, 2015; Solomon and Björk, 2011).

Asked about the amount being spent on publication fees, 22,977 SOAP survey respondents reported prices between 501 and 1000 € (12,9%), and between 1,001 and 3,000 € (9.9 %). However, more than half of the authors answering this question had not to pay a publication fee to get publish in an open access journal, and 14 % did not report the amount spent. While examining the indexing coverage of fully open access journals in bibliographic databases in 2009, Walters and Linvill (2011) also provided publication fees as journal characteristics. Their study reported an average publication fee of 923 \$, and for Journals indexed in the Web of Science 1,171 \$ on average. Solomon and Björk (2012) calculated an similar average of 906 \$ for publishing in fully open access journals. However, they found a large price variance across journals and articles. Accordingly, they suggested to cluster fully open access journals using publication fees into several groups. In this regard, while one set consisted of costly high-impact journals (2,000 \$ - 4,000 \$), a large number of journals charged 1,500 \$ - 2,000 \$. Journals with broad scope like PLOS ONE formed another segment charging between 1,000 - 1,500 \$, just as mid-priced journals from commercial publishers (500 - 1,000 \$), society journals (below 500 \$), and low-priced journals (below 200 \$).

Complementary to survey data or the consultation of journal web pages for determining spending on publication fees, some research funders and research performing institutions have recently begun to disclose their spending on publication fees as open data. To our knowledge, the first research funders providing such data were the Wellcome Trust (???) and the Austrian Science Fund FWF (???). The not-for-profit company Jisc followed by collecting data from higher-education institutions in the UK (Lawson, 2015b). Disclosed as publicly available spreadsheets, these data-sets self-report payments made along with bibliographic information, including title, journal and publisher, and a persistent identifier to the publisher's version. Curatorial efforts focused on the disambiguation of publisher and journal titles as well as on detecting duplicates. Parts of Jisc cost data was examined by Pinfield et al. (2015). Although the average spending on publication fees remained stable across the universities, they found large price variances, as well as a varying number of articles UK universities supported between 2007 - 2014.

## Central funding for publication fees in Germany

This paper focuses on how much German universities and research organization spent on open access publication fees. In Germany, the Deutsche Forschungsgemeinschaft (DFG), the largest German research funder, has strongly influenced through its “Open-Access Publishing” program how universities manage institutional support for publication fees.<sup>2</sup> Before the DFG started to pay for centrally funded publication fees on a pro rata basis in 2011, and similar to the situation described in Canada (Hampson, 2014) or the UK (Pinfield and Middleton, 2012), only few central funds existed (Eppelin et al., 2012). The DFG has enforced a set of criteria grantees have to comply with, leading to similar implementations for supporting open access publishing across German universities: these criteria exclude sponsoring of articles in hybrid journals, and the funding of articles whose publication fee exceeds 2,000 € (excluding value added tax). Grantees agree not only to reimburse the bills on behalf of the researchers they support, but also to look for ways to improve the handling of those financial transactions. They include central invoicing schemes and related agreements between university libraries and publishers. (Fournier and Weihberg, 2013).

<sup>2</sup>Guidelines for the funding program can be found here: [http://www.dfg.de/formulare/12\\_20/](http://www.dfg.de/formulare/12_20/)

Research institutes organized in the Fraunhofer-Gesellschaft, Helmholtz-Gemeinschaft, Leibniz-Gemeinschaft, and Max-Planck-Gesellschaft are not eligible for this funding program, contributing to the diversity of schemes in Germany. But in response, some organizations have adopted similar processes to support authors. The Max-Planck-Gesellschaft operates their long-lasting open access activities, including handling spending and publisher agreements centrally, through the Max Planck Digital Library (Schimmer et al., 2013; Sikora and Geschuhn, 2015), while the Leibniz-Gemeinschaft set up a dedicated open access fund in 2016.

The evolving institutional support structures to cover open access publication fees has led to calls for an unified approach towards supporting open access journal publishing. The Allianz der Wissenschaftsorganisationen<sup>3</sup>, a science policy board representing all major research organizations in Germany, marks cost transparency as one way to sustain an “adequate open access publication system” (Bruch et al., 2015). Reflecting Austrian and UK initiatives to share institutional spending on open access publication fees as open data, as well as professional discussions on open access publishing, Bielefeld University Library began to openly share its payments in May 2014. After engaging with the working group “Electronic Publishing” of the Deutsche Initiative für Netzwerkinformation (DINI)<sup>4</sup> other German institutions joined under the umbrella of the Open APC initiative soon after.

## Research question

The aim of the study was to examine how much German universities and research organisations spent on open access publication fees until 2015. Using self-reported cost data from the Open APC initiative, the analysis focused on the amount of institutional payments for publication fees, and compared these findings with those from related Austrian and UK initiatives. We also asked how thoroughly self-reported articles were indexed in CrossRef, a DOI minting agency for scholarly literature, and analysed how institutional spending per articles was distributed over publishers and journal titles.

## METHODS AND MATERIALS

We analysed self-reported cost data released by the Open APC initiative on May 13, 2016, to assess institutional spending on open access publication fees in Germany. In addition to administrative data about the amount paid per article including value added tax, the reporting institution, and the year of invoicing, we used information about whether an article was published in a fully or hybrid open access journal as well as the recorded DOI from the data-set.

We fetched bibliographic metadata for each article from CrossRef on May 19, 2016, on the basis of the reported DOIs. Although the Open APC initiative gathered metadata representing publishers and journals from CrossRef as well, this information was retrieved at the time when the participating institutions submitted the cost information. The Open APC initiative tracked the date of data submission and the contributors for every data submission with Git, a version control system, increasingly used for enabling reproducible research (Ram, 2013), and made this information available via GitHub to be transparent over time. However, CrossRef regularly updated metadata to represent ongoing mergers of publishing houses or name changes. A prominent example in this regard was the merger of the two large publishing houses Springer Business + Media and Nature Publishing Group announced on May 6, 2015, that operated as Springer Nature at the time of our study. To reflect these dynamics in academic publishing, we decided to retrieve updated metadata from CrossRef for the whole Open APC data-set instead of re-using publisher and journal information contained in the Open APC data-set.

As a client, we used the R package rcrossref (Chamberlain et al., 2016), developed and maintained by the rOpenSci initiative<sup>5</sup>, to access CrossRef’s REST API. We requested the XML-based format `application/vnd.crossref.unixsd+xml` in which full and abbreviated journal titles as well as the ISSN media types, the International Standard Serial Number used to identify journals, were distinguished. It also contained normalised publisher information, thus avoiding confusion about naming of publishing houses other studies were faced with when working with self-reported data (Woodward and Henderson, 2014). In cases where no bibliographic information could be obtained, we used the Open APC values. Because CrossRef was not the only registration agency for DOIs, but also the agencies

<sup>3</sup>[http://www.dfg.de/en/dfg\\_profile/alliance/index.html](http://www.dfg.de/en/dfg_profile/alliance/index.html)

<sup>4</sup><http://dini.de/english/ag0/e-pub0/>

<sup>5</sup>rOpenSci: <https://ropensci.org/>

144 DataCite and Medra minted DOIs for scholarly work, we furthermore obtained the DOI agency for each  
145 article with the help of the rcrossref client.

146 Data collection also involved obtaining cost data from related open data initiatives. To compare  
147 self-reported spending on open access journal articles by Germany universities and research organisations  
148 with that of other initiatives, we reviewed the openly available data-sets from the the Austrian Science  
149 Fund FWF (Reckling and Rieck, 2015; Rieck et al., 2016), Jisc (Lawson, 2015a, 2016) and the Wellcome  
150 Trust (Kiley, 2015, 2016). For analysis, we obtained the overall publication fee spending to support  
151 fully and hybrid open access journal articles. In the case of FWF, we gathered the cost information from  
152 the accompanying spending reports. We used the spreadsheet data to calculate Wellcome Trust's and  
153 Jisc's spending, and converted the prices from GBP to Euro in accordance with the average Euro foreign  
154 exchange reference rates provided by the European Central Bank. Our comparison between the open  
155 data initiatives focussed on the last two years 2014 and 2015. Because Wellcome Trust's spending was  
156 reported for the periods 2013 - 2014 and 2014 - 2015, we referred to the average exchange rates of the  
157 full two-year period as we could not determine the actual invoicing dates from the data. We excluded  
158 articles from the analysis for which neither price information nor the journal type could be obtained. In  
159 the case of Jisc's 2014 data (Lawson, 2015a), 4.861.772 € being spent on 2,812 publications were not  
160 eligible for further analysis because no publication type was given in the data-set.

161 Data collection methods of the Open APC initiative and those of the others differed in some aspects.  
162 For instance, whereas the DOI was a mandatory element in the Open APC data template that the  
163 participating institution were required to report, publication identifier in the Wellcome Trust data was  
164 also added through automated compliance checks. Our first screening of the data-sets revealed that some  
165 articles lacked a DOI. For this reason, and as our main focus is institutional funding for publication fees  
166 in Germany, we decided only to compare German spending with that of other initiatives, but not its size  
167 and distribution over publisher and journal titles, as well as the indexing coverage in CrossRef.

## 168 RESULTS

### 169 Cost Data

170 On May 13, 2016,<sup>6</sup> the Open APC initiative covered 7,417 articles whose publication fees were centrally  
171 paid by 30 German universities and research institutions until 2015. The number of supported open access  
172 journal articles grew over the years (see Figure 1). While one institution disclosed 5 payments made  
173 in 2005, the majority shared their expenditures from 2013 onwards. With 1,999 articles, the year 2015  
174 was best represented in our data-set. However, 27 institutions contributed their cost data for 2015 at the  
175 time of this analysis, suggesting that there exists a time lag between payments made and reporting these  
176 spending to the Open APC initiative.

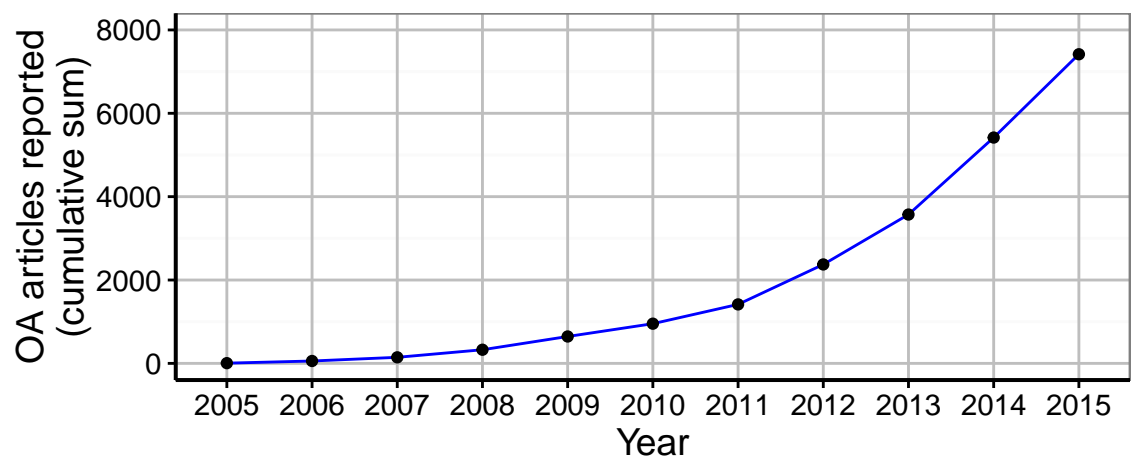
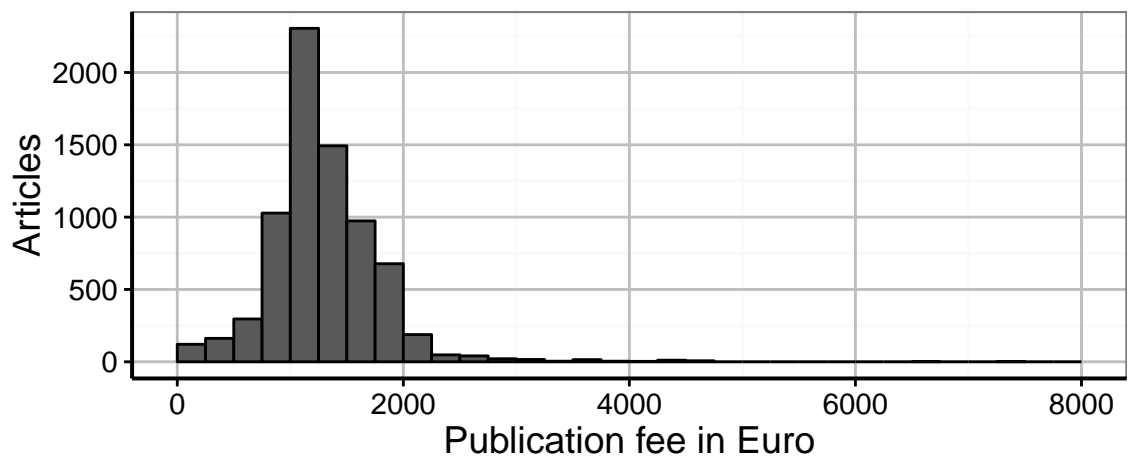


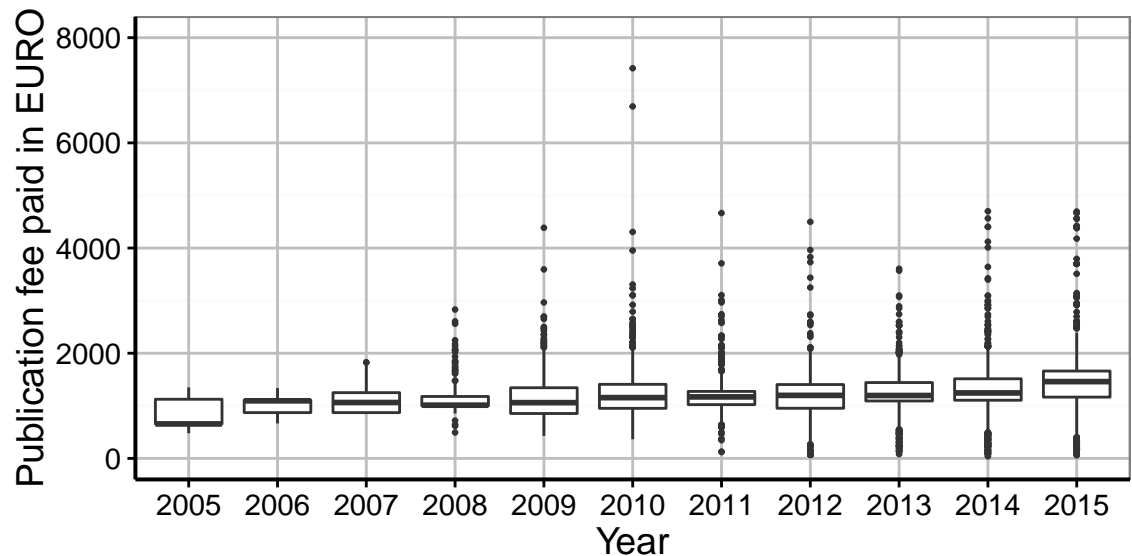
Figure 1. Growth of Open APC Initiative

<sup>6</sup>The data is openly available on GitHub. The following analysis is based on version 2.1.13 of the dataset, available at <https://github.com/OpenAPC/openapc-de/tree/v2.1.13>.



**Figure 2.** Histogram APC

177 Among all articles, fees amounted to 9,627,537 € including VAT, the average payment was 1,298 €  
 178 and the median value 1,231 €. Figure 2 shows the large price variation among the articles. The disclosed  
 179 publication fees ranged from 40 € to 7,419 €. However, the average price paid varied somewhat during  
 180 the period 2011 and 2014 (1239 - 1289 €). We also observe that 6,996 (94%) of the publication fees were  
 181 paid in accordance with the DFG price cap of 2,000 €.



**Figure 3.** Payments per year

182 The number of APC payments per institutions varied considerably (see Table 3). With 2,856 reported  
 183 articles, the Max Planck Society contributed 39 % of the overall submissions. In contrast, the two  
 184 universities of technology, TU Clausthal and TU Ilmenau, who recently begun to set up support structures  
 185 for fee-based open access journal articles, shared payments made for four articles each.

**Table 1.** Institutions self-reporting expenditures for open access publications (in €)

Institutions	Articles funded	Total	Mean	Median	Minimum	Maximum
MPG	2,856	3,661,120	1,282	1,168	69	7,419
Goettingen U	650	883,918	1,360	1,354	180	4,695
KIT	426	523,166	1,228	1,243	69	3,731

Institutions	Articles funded	Total	Mean	Median	Minimum	Maximum
Regensburg U	399	503,205	1,261	1,207	77	4,403
Muenchen LMU	365	463,491	1,270	1,299	496	2,023
TU Muenchen	308	390,086	1,267	1,386	131	2,122
Bielefeld U	262	322,815	1,232	1,234	142	2,103
Giessen U	243	326,082	1,342	1,247	81	4,498
Konstanz U	221	302,659	1,369	1,380	40	2,072
Heidelberg U	215	308,348	1,434	1,500	60	2,042
Wuerzburg U	207	286,543	1,384	1,447	105	2,514
Leipzig U	173	243,873	1,410	1,471	341	2,055
FZJ - ZB	158	196,869	1,246	1,177	370	3,700
TU Dresden	130	175,723	1,352	1,415	200	2,193
Duisburg-Essen U	114	136,911	1,201	1,214	238	1,982
FU Berlin	106	142,671	1,346	1,292	220	2,000
GFZ-Potsdam	106	126,520	1,194	1,065	223	4,403
Bayreuth U	92	105,725	1,149	1,200	82	2,059
Bochum U	71	93,546	1,318	1,438	100	2,042
Hannover U	69	90,259	1,308	1,241	149	2,159
MDC	69	145,256	2,105	1,800	491	4,700
TU Chemnitz	36	37,826	1,051	1,142	78	2,123
Kassel U	35	35,550	1,016	1,142	150	1,861
Hamburg TUHH	24	32,789	1,366	1,466	300	2,027
Potsdam U	24	32,128	1,339	1,386	916	2,116
Bamberg U	22	23,663	1,076	1,009	90	2,010
TU Ilmenau	13	13,053	1,004	986	178	2,077
Dortmund TU	9	8,238	915	900	155	1,738
TU Clausthal	8	6,999	875	918	181	1,724
INM - Leibniz-Institut für Neue Materialien	6	8,505	1,418	1,492	237	2,454

### Comparison of related cost data-sets

Table 2 compares Open APC spending data with that of the Austrian FWF, as well as with Jisc's and Wellcome Trust's expenditures. Prices were converted according to the average Euro exchange rate of the examined periods, and gathered for both fully and hybrid open access journals. The comparison reveals that the Open APC initiative lacked cost information about hybrid journals, whereas the related Austrian and UK open data initiatives could report a large share of spending for these journals between 2014 and 2015. This situation likely reflected the DFG-inspired funding policy of most participating institutions that excluded support for publications in hybrid journals. Over the years 2005 – 2015, 3 out of 30 German universities and research institutions reported 60 hybrid journal articles to the Open APC initiative, accounting for 0.81 % of the overall payments. In terms of the number of supported articles and the amount being spent on publication fees, by contrast, the Open APC data-set provided the most comprehensive price information for fully open access journals compared to what the Austrian and UK initiatives had reported.

**Table 2.** Comparison of cost data per period and journal type (in €)

Cost data-set	Journal Type	Articles funded	Total costs in Euro	Mean
FWF		NA	NA	NA
2014	Fully OA	247	316,765	1,282
	Hybrid OA	780	1,794,604	2,301
2015	Fully OA	288	418,408	1,453

Cost data-set	Journal Type	Articles funded	Total costs in Euro	Mean
Jisc 2014	Hybrid OA	912	2,376,356	2,606
		NA	NA	NA
	Fully OA	1,161	1,897,862	1,635
2015	Hybrid OA	2,938	5,409,623	1,841
	Fully OA	1,168	2,211,958	1,894
	Hybrid OA	2,944	6,977,753	2,370
Open APC 2014		NA	NA	NA
	Fully OA	1,832	2,353,665	1,285
	Hybrid OA	15	26,546	1,770
2015	Fully OA	1,991	2,820,445	1,417
	Hybrid OA	8	23,412	2,927
		NA	NA	NA
Wellcome Trust 2013-2014	Fully OA	607	911,302	1,501
	Hybrid OA	1,894	4,648,878	2,455
	Fully OA	775	1,418,097	2,756
2014-2015	Hybrid OA	2,065	5,690,178	1,830

Comparison of average prices suggests that publishing in hybrid journal was more expensive than in fully open access journals. Price differentials between these two categories were also reported earlier, concluding that prices for fully open access journals were generally lower (Pinfield et al., 2015; Solomon and Björk, 2012). In 2014 and 2015, the mean price for fully open access journals calculated from all data-sets was below the DFG price cap of 2,000 €.

### CrossRef indexing

Along with the price information, participating institutions were required to identify funded articles by their DOI. They were reported for 7,373 out of 7,417 articles. Of those, `my_apc %>% filter(indexed.in.crossref == TRUE) %>% nrow()` were indexed in CrossRef, representing `my_apc %>% filter(indexed.in.crossref == TRUE) %>% nrow() / my_apc %>% nrow() * 100` % of all funded publications. The reasons why articles identified by a DOI were not registered with CrossRef differed. Some journals were not indexed by CrossRef at the time of our study but by the DOI agencies DataCite (Journal of new frontiers in spatial concepts published by KIT Scientific Publishing) and Medra (DIE ERDE: Journal of the Geographical Society of Berlin). In other cases, either the DOI did not refer to the full text despite the fact that the journal was indexed on a regular basis (compare <http://doi.org/10.1186/1471-2105-13-S19-S7> with <http://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-13-S19-S7>) or the resource type could not be retrieved, although the DOI resolves (<http://doi.org/10.1186/s12885-015-1795-7>).

### Cost data by publisher and journal

We used the DOI to automatically fetch publisher and journal names for each article from the CrossRef REST API. Table 4 shows the top ten publishers in terms of payments made that represent 92 % of the spending for publication fees. In total, payments were made to 139 publishing houses. In comparison with data from the UK, full open access publishers have a greater share on total spending. Pinfield et al. (2015), for instance, reported remarkably lower numbers for the open access publishers MPDI AG, Copernicus GmbH, and Hindawi Publishing.

**Table 3.** Publication fees paid per publisher (in €)

Journal	Articles funded	Total	Mean	SD	Median	Range
Springer Nature	2,167	2,948,697	1,361	387	1,385	80.92 – 4403

Journal	Articles funded	Total	Mean	SD	Median	Range
Public Library of Science (PLOS)	1,680	2,243,128	1,335	321	1,207	555.66 – 2790.27
Frontiers Media SA	906	1,186,283	1,309	424	1,142	77.35 – 4179
Copernicus GmbH	841	1,160,450	1,380	658	1,277	69.12 – 7418.88
IOP Publishing	677	699,137	1,033	228	953	374.77 – 1950
MDPI AG	208	236,729	1,138	453	1,177	154.43 – 2054.68
Hindawi Publishing Corporation	120	125,495	1,046	538	947	174.99 – 2225.22
The Optical Society	111	176,665	1,592	392	1,626	498.62 – 3731.09
Wiley-Blackwell	78	126,148	1,617	467	1,601	490.58 – 3065
Oxford University Press (OUP)	64	118,225	1,847	793	1,741	297.5 – 4498.2
Other	565	606,578	1,074	840	922	40 – 4699.61

Most of the funding of publication fees in Germany went to the publisher Springer Science + Business Media, especially profiting from the merge with the former full open access publisher BioMed Central. In contrast, other established publishing houses such as Elsevier and Wiley-Blackwell rank lower, presumably because they mostly publish hybrid journals, which were not well represented in our data-set at the time of the study. Table 4 also illustrates the variation across and within publishers, which confirms earlier findings (Pinfield et al., 2015).

**Table 4.** Publication fees paid per journal (in €)

Journal	Articles funded	Total	Mean	SD	Median	Range
PLOS ONE	1,433	1,745,513	1,218	130	1,198	748.71 – 1808.8
New Journal of Physics	673	693,322	1,030	225	953	374.77 – 1856.4
Atmospheric Chemistry and Physics Discussions	281	437,903	1,558	776	1,403	233.86 – 7418.88
Frontiers in Psychology	271	363,794	1,342	429	1,142	77.35 – 2122.81
BMC Genomics	135	179,592	1,330	205	1,276	920 – 1926
Biogeosciences Discussions	127	187,716	1,478	548	1,313	663.55 – 3641.47
BMC Bioinformatics	113	142,680	1,263	217	1,244	655 – 1661.24
Frontiers in Plant Science	107	126,763	1,185	408	1,106	551.04 – 2380



Journal	Articles funded	Total	Mean	SD	Median	Range
Atmospheric Measurement Techniques Discussions	107	143,782	1,344	585	1,203	428.4 – 3709.44
Frontiers in Human Neuroscience	106	140,065	1,321	415	1,106	575 – 2000
Other	4,064	5,466,407	1,345	557	1,350	40 – 4699.61

Prices also varied within single journals. Based on the number of articles paid for, Table 5 illustrates the top ten out of 732 journals. Payments to these ten journals represent 45 % of all payments. In the case of Atmospheric Chemistry and Physics Discussions, the price range can be explained by the fact that this journal charges per page and also takes the submission's file format into consideration.

## DISCUSSION

In Germany, institutional spending on open access publication fees has increased over the years, confirming the general trend of publication fees in academic publishing (Davis and Walters, 2011; Laakso and Björk, 2012; Pinfield, 2015). With a share of 99%, the majority of open access articles German institutions reported to the Open APC initiative were published in fully open access journals. This presumably reflects the DFG funding policy which excludes the support of articles published in hybrid open access journals. The DFG has been financially supporting the implementation of central publication funds at more than 30 German universities since 2009. However, in reviewing self-reported cost data from funders or countries that also support hybrid open access journals or open access books, we revealed smaller proportions of payments in favour of articles in fully open access journals. Because open access publication fee spending is fragmented, we cannot answer whether German researchers avoided opting for open access when publishing in hybrid journals or used other budgets to pay publication fees required to make their work open access through these kind of journals.

In our study, CrossRef thoroughly indexed open access journal articles in the Open APC data-set. We could gather metadata representing publisher and journal titles for 99 % articles, and successfully merged these information with the Open APC cost data. Using metadata from CrossRef, therefore, reduces the extensive validation work of bibliographic information provided that the reporting of the DOI along with the expenses is made mandatory. Drawing on CrossRef would also increase the comparability of cost data for future negotiations with publishers on open access agreements, and the open access spending between open data initiatives that apply the same reporting standards, as its metadata represent the dynamic landscape of academic publishing in terms of ongoing mergers of publishing houses or name changes.

Another advantage of self-reported data-sets on the article-level to disclose spending on open access publication fees is that they enable researchers and practitioners alike to study in which open access journals researchers from one institution actually publish, and to compare these findings with that of other universities or research organizations. For instance, our study revealed that the size of publication fee spending differed among the institutions with the Max Planck Society accounting for almost 39 % of the overall articles. Many universities and research organization reported remarkably lower number of supported open access articles to the Open APC initiative. Using self-reported data, therefore, contributes to the understanding about how much and to what extent spending on open access publishing varies on the institutional level. This is particularly relevant given the increasingly important role open access publishing plays in recent negotiations between German universities and research organizations forming consortia on the one side and publishers on the other about financing scholarly publishing in future (Bruch et al., 2015).

This study is limited in some respects. One is that we cannot assess whether publishers and journals granted publication fee discounts seeing that the Open APC initiative does not track this kind of information. However, the large price ranges of particular journals suggests that varying pricing schemes were in place. Adding to this complexity, it is likely that some institutions only paid parts of the publication fee.

Take for instance the journal Nature Communication. Charges reported in our sample ranged between 2000 €, the DFG price cap, and 4.403 €. Such co-payments that involve several budgets were a proposed strategy to sustain publication funds at German universities (Fournier and Weihberg, 2013). In another case, one university included its charges for participating in the German SCOAP consortia and presumably divided the sum by the articles published in SCOAP journals. In this national consortia, managed by the German National Library of Science and Technology, payments were not directly made per article. Instead, subscription costs between a participating library and a publisher were reconciled, and the reduction transferred to the consortia to finance publications in SCOAP journals.

It must also be noted that reporting to the Open APC initiative is voluntary. Therefore, not all institutions in Germany that provide central funding of publication fees contribute cost data to this initiative. In a qualitative survey, asking why German institutions are reluctant to share their cost data through the Open APC initiative, one institution feared that increase in transparency would allow publishers to adjust prices in their favour. Others pointed out that the workload to produce such a data-set could be too extensive (Deppe, 2015). As no registry of institutional open access funds or similar support structures exists, we cannot assess the number of non-participants in Germany.

Our analysis on how institutional spending per articles was distributed over publishers and journal titles shows that open access publishing is diverse and concentrated at the same time. While we were able to identify 139 individual publishing houses that were supported by the German universities and research organizations, the distribution is highly skewed. 92 % of open access publication fee spending went to ten publishers, confirming a general high concentration of few publishers in current academic publishing. However, our study could not confirm that publications in open access journals owned by traditional publishing houses account for most of the spending on publication fees. Rather, open access publishers such as Public Library of Science (PloS), Copernicus GmbH or MPDI AG rank higher in our study than in the analyses of cost data in the UK.

This study finally confirms the leading role of “mega-journals” in open access publishing, including the multidisciplinary PLOS ONE and the journals New Journal of Physics, Atmospheric Chemistry and Physics Discussions and Frontiers in Psychology, all of which publish contributions from all branches of their respective discipline. In general, an estimated 14 out of more than 10,000 journals registered in DOAJ in 2015 accounted for up to 15–20 % of all articles published in full open access journals (Björk, 2015).

## CONCLUSION

## ACKNOWLEDGMENT

We thank Andrea Hacker and Ada-Charlotte Regelmann for valuable comments on the first draft of this paper. We also thank Christoph Broschinski, Vitali Peil, and Dirk Pieper, the members of the DINI working group “Electronic publishing”, and all data contributors<sup>7</sup> of the Open APC initiative.

## REFERENCES

- Björk, B.-C. (2015). Have the ‘mega-journals’ reached the limits to growth? *PeerJ* 3, e981. <http://doi.org/10.7717/peerj.981>.
- Björk, B.-C., and Solomon, D. (2012). Pricing principles used by scholarly open access publishers. *Learned Publishing* 25, 132–137. <http://doi.org/10.1087/20120207>.
- Björk, B.-C., and Solomon, D. (2014). How research funders can finance APCs in full OA and hybrid journals. *Learned Publishing* 27, 93–103. <http://doi.org/10.1087/20140203>.
- Björk, B.-C., and Solomon, D. (2015). Article processing charges in OA journals: relationship between price and quality. *Scientometrics* 103, 373–385. <http://doi.org/10.1007/s11192-015-1556-z>.
- Bruch, C., Deinzer, G., Geschuhn, K., Haetscher, P., Hillenkoetter, K., Kress, U., et al. (2015). Positions on creating an Open Access publication market which is scholarly adequate : Positions of the Ad Hoc Working Group Open Access Gold in the priority initiative “Digital Information” of the Alliance of Science Organisations in Germany. Ad-hoc-Arbeitsgruppe Open-Access-Gold der Schwerpunktinitiative

<sup>7</sup><https://github.com/OpenAPC/openapc-de#contributors>

320 “Digitale Information” der Allianz der deutschen Wissenschaftsorganisationen. <http://doi.org/10.2312/allianzoa.009>.

321

322 Chamberlain, S., Boettiger, C., Hart, T., and Ram, K. (2016). *rcrossref: Client for Various ‘CrossRef’*

323 *APIs*. Available at: <https://CRAN.R-project.org/package=rcrossref>.

324 Dallmeier-Tiessen, S., Darby, R., Goerner, B., Hyppoelae, J., Igo-Kemenes, P., Kahn, D., et al. (2011).

325 Highlights from the SOAP project survey. What Scientists Think about Open Access Publishing. Available

326 at: <http://arxiv.org/abs/1101.5260>.

327 Davis, P. M., and Walters, W. H. (2011). The impact of free access to the scientific literature: a review

328 of recent research. *Journal of the Medical Library Association* 99, 208–217. <http://doi.org/10.3163/1536-5050.99.3.008>.

329

330 Deppe, A. (2015). *Ansätze zur Verstetigung von Open-Access-Publikationsfonds.*, ed. K. Umlauf

331 Institut für Bibliotheks- und Informationswissenschaft. Available at: <http://nbn-resolving.de/urn:nbn:de:kobv:11-100234262>.

332

333 Eppelin, A., Pampel, H., Bandilla, W., and Kaczmirek, L. (2012). Umgang mit Open-Access-

334 Publikationsgebühren – die Situation in Deutschland in 2010. *GMS Medizin — Bibliothek — Information*.

335 12. <http://doi.org/10.3205/mbi000240>.

336 Fournier, J., and Weihberg, R. (2013). Das Förderprogramm “Open Access Publizieren” der Deutschen

337 Forschungsgemeinschaft. Zum Aufbau von Publikationsfonds an wissenschaftlichen Hochschulen in

338 Deutschland. *Zeitschrift für Bibliothekswesen und Bibliographie* 60, 236–243. <http://doi.org/10.3196/186429501360528>.

339

340 Gumpenberger, C., Ovalle-Perandones, M.-A., and Gorraiz, J. (2012). On the impact of Gold Open

341 Access journals. *Scientometrics* 96, 221–238. [http://doi.org/10.1007/s11192-012-0902-](http://doi.org/10.1007/s11192-012-0902-7)

342 [7](http://doi.org/10.1007/s11192-012-0902-7).

343 Hampson, C. (2014). The Adoption of Open Access Funds Among Canadian Academic Research

344 Libraries, 2008-2012. *Partnership: The Canadian Journal of Library and Information Practice and*

345 *Research* 9. <http://doi.org/10.21083/partnership.v9i2.3115>.

346 Kiley, R. (2015). Wellcome Trust open access (OA) spend and compliance monitoring: 2013-14.

347 <http://doi.org/10.6084/m9.figshare.1321361.v5>.

348 Kiley, R. (2016). Wellcome Tust/COAF spend on open access publishing (article processing charges)

349 - 2014-15. <http://doi.org/10.6084/m9.figshare.3118936.v1>.

350 Kingsley, D. A. (2014). Paying for Publication: Issues and Challenges for Research Support Services.

351 *Australian Academic & Research Libraries* 45, 262–281. [http://doi.org/10.1080/00048623.](http://doi.org/10.1080/00048623.2014.945135)

352 [2014.945135](http://doi.org/10.1080/00048623.2014.945135).

353 Laakso, M., and Björk, B.-C. (2012). Anatomy of open access publishing: a study of longitudinal

354 development and internal structure. *BMC Medicine* 10, 124. [http://doi.org/10.1186/1741-](http://doi.org/10.1186/1741-7015-10-124)

355 [7015-10-124](http://doi.org/10.1186/1741-7015-10-124).

356 Lawson, S. (2015a). APC data for 25 UK higher education institutions - 2014. <http://doi.org/10.6084/m9.figshare.1305596.v5>.

357

358 Lawson, S. (2015b). Article Processing Charges Paid by 25 UK Universities in 2014. *Journal of Open*

359 *Humanities Data* 1. <http://doi.org/10.5334/johd.2>.

360 Lawson, S. (2015c). Fee Waivers for Open Access Journals. *Publications* 3, 155–167. <http://doi.org/10.3390/publications3030155>.

361

362 Lawson, S. (2016). APC data for 27 UK higher education institutions in 2015. <http://doi.org/10.6084/m9.figshare.1507481.v4>.

363

364 Lawson, S., and Meghreblian, B. (2015). Journal subscription expenditure of UK higher education

365 institutions. *F1000Research*. <http://doi.org/10.12688/f1000research.5706.3>.

366 Noorden, R. V. (2013). Open access: The true cost of science publishing. *Nature* 495, 426–429.

367 <http://doi.org/10.1038/495426a>.

368 Pinfield, S. (2015). Making Open Access work. *Online Information Review* 39, 604–636. <http://doi.org/10.1108/oir-05-2015-0167>.

369

370 Pinfield, S., and Middleton, C. (2012). Open access central funds in UK universities. *Learned*

371 *Publishing* 25, 107–117. <http://doi.org/10.1087/20120205>.

372 Pinfield, S., Salter, J., and Bath, P. A. (2015). The “total cost of publication” in a hybrid open-access

373 environment: Institutional approaches to funding journal article-processing charges in combination with

374 subscriptions. *Journal of the Association for Information Science and Technology*. [http://doi.org/](http://doi.org/10.1002/asi.23446)  
375 [10.1002/asi.23446](http://doi.org/10.1002/asi.23446).

376 Ram, K. (2013). Git can facilitate greater reproducibility and increased transparency in science.  
377 *Source Code for Biology and Medicine* 8, 7. <http://doi.org/10.1186/1751-0473-8-7>.

378 Reckling, F., and Rieck, K. (2015). Austrian Science Fund (FWF) Publication Cost Data 2014.  
379 <http://doi.org/10.6084/m9.figshare.1378610.v14>.

380 Rieck, K., Haslinger, D., Meischke-Ilic, S., Kirindi-Hentschel, Ü., and Reckling, F. (2016). Austrian  
381 Science Fund (FWF) Publication Cost Data 2015. [http://doi.org/10.6084/m9.figshare.](http://doi.org/10.6084/m9.figshare.3180166.v1)  
382 [3180166.v1](http://doi.org/10.6084/m9.figshare.3180166.v1).

383 Schimmer, R., Geschuhn, K., and Palzenberger, M. (2013). Open Access in Zahlen: Der Umbruch in  
384 der Wissenschaftskommunikation als Herausforderung für Bibliotheken. *Zeitschrift für Bibliothekswesen*  
385 *und Bibliographie* 60, 244–250. <http://doi.org/10.3196/186429501360532>.

386 Sikora, A., and Geschuhn, K. K. (2015). Management of article processing charges – challenges for  
387 libraries. *Insights: the UKSG journal* 28, 87–92. <http://doi.org/10.1629/uksg.229>.

388 Solomon, D. J., and Björk, B.-C. (2011). Publication fees in open access publishing: Sources of  
389 funding and factors influencing choice of journal. *Journal of the Association for Information Science and*  
390 *Technology* 63, 98–107. <http://doi.org/10.1002/asi.21660>.

391 Solomon, D. J., and Björk, B.-C. (2012). A study of open access journals using article processing  
392 charges. *Journal of the Association for Information Science and Technology* 63, 1485–1495. <http://doi.org/10.1002/asi.22673>.

393 Suber, P. (2012). *Open Access*. MIT Press. Available at: [https://mitpress.mit.edu/](https://mitpress.mit.edu/books/open-access)  
394 [books/open-access](https://mitpress.mit.edu/books/open-access).

395 Walters, W. H., and Linvill, A. C. (2011). Bibliographic index coverage of open-access journals in  
396 six subject areas. *Journal of the Association for Information Science and Technology* 62, 1614–1628.  
397 <http://doi.org/10.1002/asi.21569>.

398 Woodward, H. M., and Henderson, H. L. (2014). Report for Jisc Collections on total cost of owner-  
399 ship project: Data capture and process. Information Power Ltd. Available at: [https://www.jisc-](https://www.jisc-collections.ac.uk/Global/News%20files%20and%20docs/IPL-Jisc-Total-Cost-of-Ownership-Data-Capture-Report.pdf)  
400 [collections.ac.uk/Global/News%20files%20and%20docs/IPL-Jisc-Total-Cost-](https://www.jisc-collections.ac.uk/Global/News%20files%20and%20docs/IPL-Jisc-Total-Cost-of-Ownership-Data-Capture-Report.pdf)  
401 [of-Ownership-Data-Capture-Report.pdf](https://www.jisc-collections.ac.uk/Global/News%20files%20and%20docs/IPL-Jisc-Total-Cost-of-Ownership-Data-Capture-Report.pdf).