Report from Germany & experiences with the OpenAPC workflow

Open Access Monitoring – Approaches and Perspectives
University of Vienna | April 09-10, 2018

Dirk Pieper (Bielefeld UL)

Agenda

1. **INTACT**

2. OpenAPC approach

3. OpenAPC data analysis

4. Open Access monitoring in Germany

INTACT („transparent infrastructure for open access publication fees") funded by German Research Foundation (DFG) for 3 years (until 09/2018)

Partners: Bielefeld UL, Max Planck Digital Library, Institute for Interdisciplinary Studies of Science (I²SoS, Bielefeld U)

Supported by the DINI working group Electronic Publishing

ESAC (Lead: MPDL)

Help shaping the principles and conditions of APC based business models, network to communicate good practices and guidelines

Gain more evidence on institutional APC workflows and their respective challenges involved

Partner with publishers to work on the improvement of APC workflows

ESAC: Outcomes

➢ Workshop on the effectiveness of APCs (2018-06, Munich)

➢ Offsetting dataset, Offsetting coverage analysis , Springer Compact analysis

➢ Geschuhn, K., & Stone, G. (2017). It's the workflows, stupid! What is required to make 'offsetting' work for the open access transition. Insights, 30(3), 103–114. DOI: http://doi.org/10.1629/uksg.391

ESAC: Outcomes

➢ Offsetting workshops 2017-03 (Vienna), 2016-03 (Munich)

➢ INTACT workshop offsetting and open access transformation for german acquisition librariens (2017-05)

➢ [Customer Recommendations for Article Workflows and Services for Offsetting](#) (2017-03)

➢ [Joint understanding of offsetting](#) (2016-03)

OA analytics (Lead: I²SoS)

Provides bibliometric indicators about the development of open access publishing in universities and research organisations

➢ ISSN-GOLD-OA (most comprehensive list of OA journal titles so far, e.g. used in OpenAPC to identify hybrid journals)

➢ Publications in gold oa journals on a global and european scale (only in german language so far)

![INT@CT - Transparent Infrastructure for Article Charges / open APC | ESAC | OA analytics]

OA percentages for countries with the biggest publication output worldwide based on Web of Science data 2008-2016

Land	Anzahl Publikationen (2008-2016)	Anzahl Gold OA-Publikationen (2008-2016)	Gold OA-Publikationen (%)	Veränderung Anteil Gold OA (%)*	Rang
USA	3.730.737	297.216	7,97	0,89	1
China	2.028.043	233.319	11,50	1,39	2
UK	1.074.510	97.476	9,07	1,02	3
Germany	948.886	99.745	10,51	1,05	4
Japan	754.782	91.190	12,08	1,18	5
France	668.616	62.240	9,31	0,91	6
Canada	596.427	55.808	9,35	0,95	7
Italy	579.223	63.487	10,96	0,99	8
Spain	504.908	66.630	13,19	0,76	9
India	492.111	65.541	13,32	-0,23	10

* Angegeben wird hier die durchschnittliche jährliche Veränderung des Gold-OA-Anteils an den Gesamtpublikationen in Prozent.

Wohlgemuth M, Rimmert C, Taubert NC. *Publikationen in Gold-Open-Access-Journalen auf globaler und europäischer Ebene sowie in Forschungsorganisationen*. Bielefeld: Universität Bielefeld, https://pub.uni-bielefeld.de/publication/2912807, Table 2.1

Countries with the highest oa percentages worldwide based on Web of Science data 2008-2016

Land	Anzahl Publikationen (2008-2016)	Anzahl Gold OA-Publikationen (2008-2016)	Gold OA-Publikationen (%)	Veränderung Anteil Gold OA (%)*	Rang
Brasil	347.014	117.596	33,88	-1,06	1
Serbia	42.525	12.510	29,42	-0,42	2
Pakistan	54.081	15.605	28,85	-0,81	3
Croatia	33.537	9.665	28,81	-0,60	4
Colombia	30.861	8.514	27,58	-0,71	5
Chile	57.982	15.754	27,17	-0,32	6
Kenya	12.240	3.184	26,01	1,47	7
Nigera	21.275	5.366	25,22	-1,53	8
Bangladesh	12.545	2.872	22,89	0,40	9
Slovenia	34.461	6.978	20,25	0,12	10

* Angegeben wird hier die durchschnittliche jährliche Veränderung des Gold-OA-Anteils an den Gesamtpublikationen in Prozent.

Wohlgemuth M, Rimmert C, Taubert NC. *Publikationen in Gold-Open-Access-Journalen auf globaler und europäischer Ebene sowie in Forschungsorganisationen*. Bielefeld: Universität Bielefeld, https://pub.uni-bielefeld.de/publication/2912807, Table 2.2

OA percentages for european countries with the biggest publication output based on Web of Science data 2008-2016

Land	Anzahl Publikationen (2008-2016)	Anzahl Gold OA-Publikation (2008-2016)	Gold OA-Publikationen (%)	Veränderung Anteil Gold OA (%)*	Rang
UK	1.074.510	97.476	9,07	1,02	1
Germany	948.886	99.745	10,51	1,05	2
France	668.616	62.240	9,31	0,91	3
Italy	579.223	63.487	10,96	0,99	4
Spain	504.908	66.630	13,20	0,76	5
Netherlands	334.870	36.115	10,78	1,15	6
Poland	218.331	41.739	19,12	0,36	7
Sweden	217.122	27.631	12,73	1,20	8
Belgium	186.397	19.490	10,46	1,14	9
Denmark	136.743	16.037	11,73	1,12	10

* Angegeben wird hier die durchschnittliche jährliche Veränderung des Gold-OA-Anteils an den Gesamtpublikationen in Prozent.

Wohlgemuth M, Rimmert C, Taubert NC. *Publikationen in Gold-Open-Access-Journalen auf globaler und europäischer Ebene sowie in Forschungsorganisationen*. Bielefeld: Universität Bielefeld, https://pub.uni-bielefeld.de/publication/2912807, Table 3.1

European countries with the highest oa percentages based on Web of Science data 2008-2016

Land	Anzahl Publikationen (2008-2016)	Anzahl Gold OA-Publikationen (2008-2016)	Gold OA-Publikationen (%)	Veränderung Anteil Gold OA (%)*	Rang
Croatia	33.537	9.665	28,82	-0,60	1
Slovenia	34.461	6.978	20,25	0,12	2
Estonia	15.338	3.012	19,64	0,57	3
Poland	218.331	41.739	19,12	0,36	4
Slovakai	33.850	5.617	16,59	0,03	5
Czech Republic	107.904	17.564	16,28	0,30	6
Lithuania	22.824	3.337	14,62	0,70	7
Bulgaria	22.937	3.270	14,26	1,35	8
Romania	82.595	10.963	13,27	1,03	9
Spain	504.908	66.630	13,20	0,76	10

* Angegeben wird hier die durchschnittliche jährliche Veränderung des Gold-OA-Anteils an den Gesamtpublikationen in Prozent.

Wohlgemuth M, Rimmert C, Taubert NC. *Publikationen in Gold-Open-Access-Journalen auf globaler und europäischer Ebene sowie in Forschungsorganisationen*. Bielefeld: Universität Bielefeld, https://pub.uni-bielefeld.de/publication/2912807, Table 3.2

Some results and hypotheses to be tested:

Growth rate of Gold-OA-Percentages per year between 2008-2016 is too small, to reach (national/european/global) political goals of Open Access or Open Science guidelines

Number of Gold-OA-Journals is too small, to reach higher growth rates, which affirms OA2020 approach to transform existing journals

Current Springer Compact agreements increase the number of open access articles in hybrid journals (2016-2017), but not enough to flip journals entirely

OpenAPC (Lead: Bielefeld UL)

Started by Bielefeld UL in 2014-06, goals:

- to release datasets on fees paid for Open Access journal articles by Universities and Research Society Funds under an Open Database License
- to demonstrate how reporting on fee-based Open Access publishing can be made more transparent and reproducible across institutions
- to provide valid APC cost data for Open Access transformation

Agenda

1. INTACT

2. **OpenAPC approach**

3. OpenAPC data analysis

4. Open Access monitoring in Germany

OpenAPC approach

Require as less as information as possible from contributing institutions, see [Data Submission Handout](#) and [OpenAPC schema](#)

Use Crossref API to normalise bibliographic metadata (via DOI)

Enrich identifiers (ISSNs, PMID, PMCID, UT)

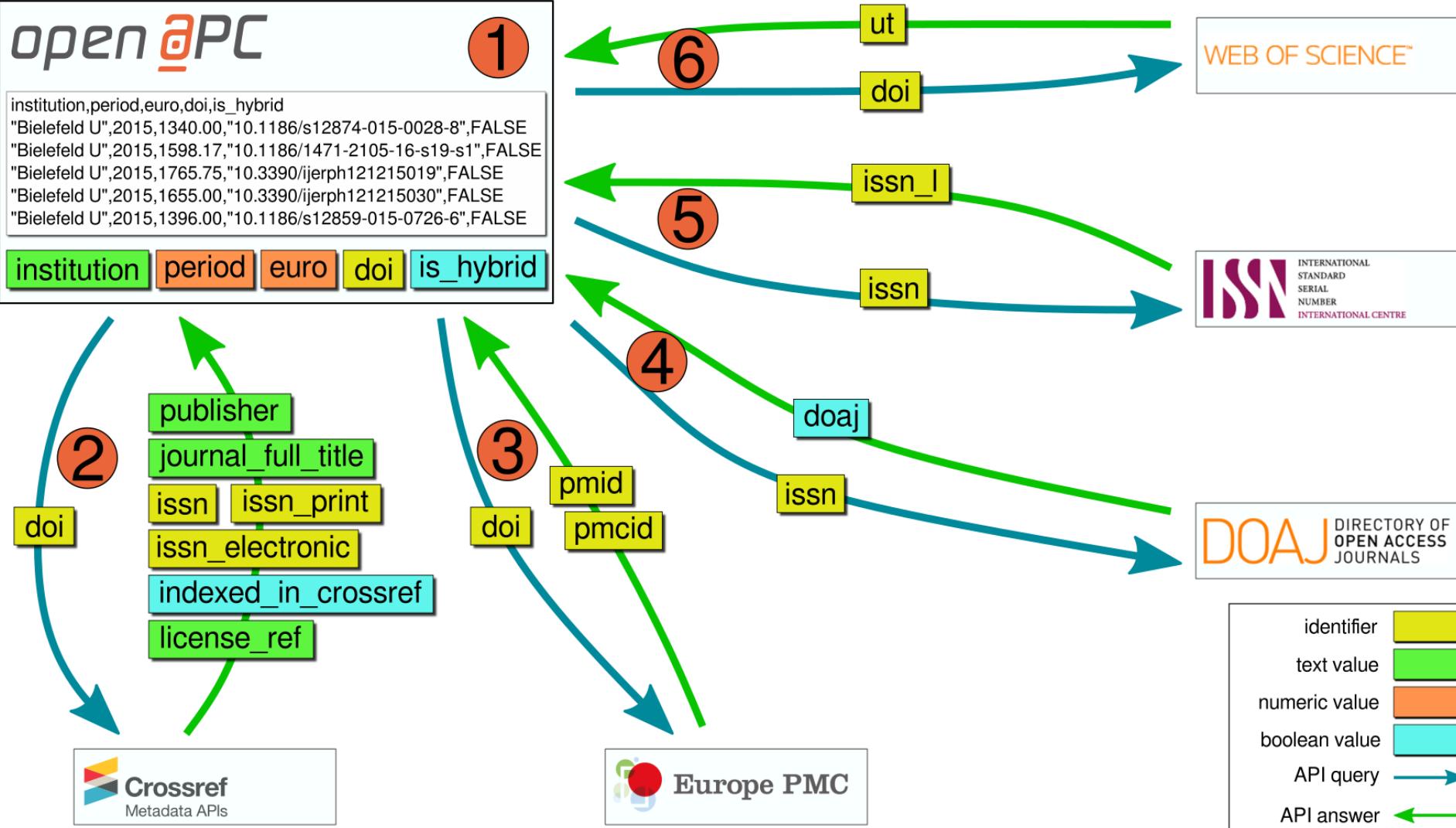Continous [data integrity testing](#)

OpenAPC approach

Mandatory fields (https://github.com/OpenAPC/openapc-de/wiki/Data-Submission-Handout)

➢ Institution: Top-level organisation which covered the fee
➢ Period: Year of APC payment
➢ The final amount that was paid, incl. VAT and all additional fees (special tax rates, prepayment discounts, central billing agreements or individual waivers can be documented, currencies can be converted using APIs of European Central Bank)

OpenAPC approach

Mandatory fields (https://github.com/OpenAPC/openapc-de/wiki/Data-Submission-Handout)

➢ DOI: Digital Object Identifier
➢ is_hybrid: Should be TRUE if the article was published in a subscription-based Journal ('hybrid journal'), FALSE if the journal was fully Open Access

OpenAPC enrichment overview

OpenAPC approach

Scripts for enrichment steps, normalisation, quality checks and analysis are open source

Version control through GitHub including history (automatic snyc to local GitLab)

No separation of data, documentation and reporting code

Data in common format (CSV, JSON), plus OLAP Server API

OpenAPC approach

Blog, calculations, plots automatically created using (R) markdown

Visualisation layer (treemaps) through Cubes OLAP Server

Datasets are made available under open data commons license v1.0

OpenAPC meets most of the FAIR data principles

Agenda

1. INTACT
2. OpenAPC approach
3. **OpenAPC data analysis**
4. Economic relevance of OpenAPC

INTaCT _Transparent Infrastructure for Article Charges_

open aPC ESaC Oa analytics

Latest release

v3.12.3
d3738c9

FU Berlin APC data 2016

cbroschinski released this 22 hours ago · 5 commits to master since this release

Contribution from FU Berlin contains remaining APC data for the 2016 period.

Downloads

Source code (zip)

Source code (tar.gz)
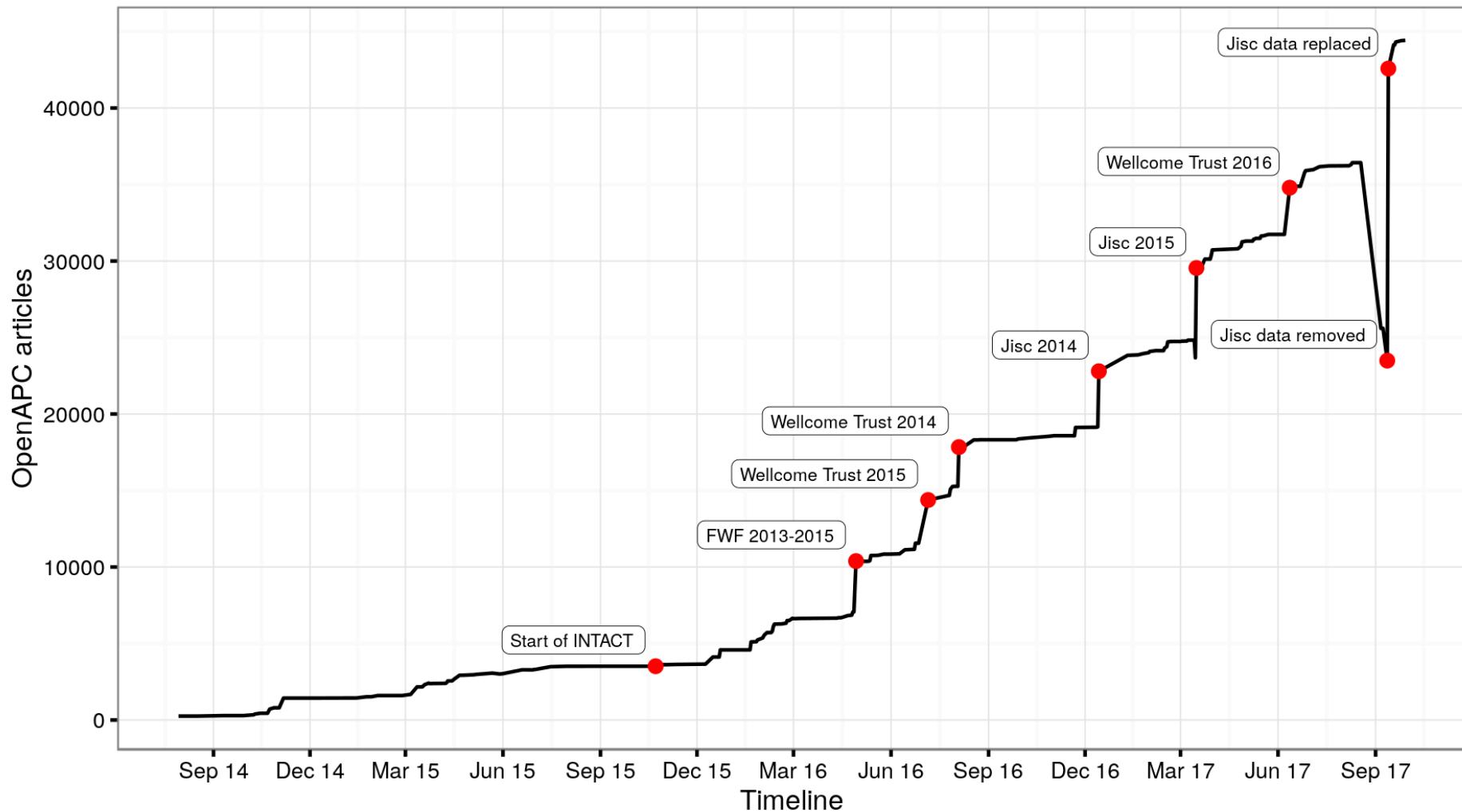
v3.12.2
d7583e9

Oldenburg U APC data 2017

cbroschinski released this 6 days ago · 7 commits to master since this release

Contribution from Oldenburg University contains APC data for the 2016 and 2017 periods.
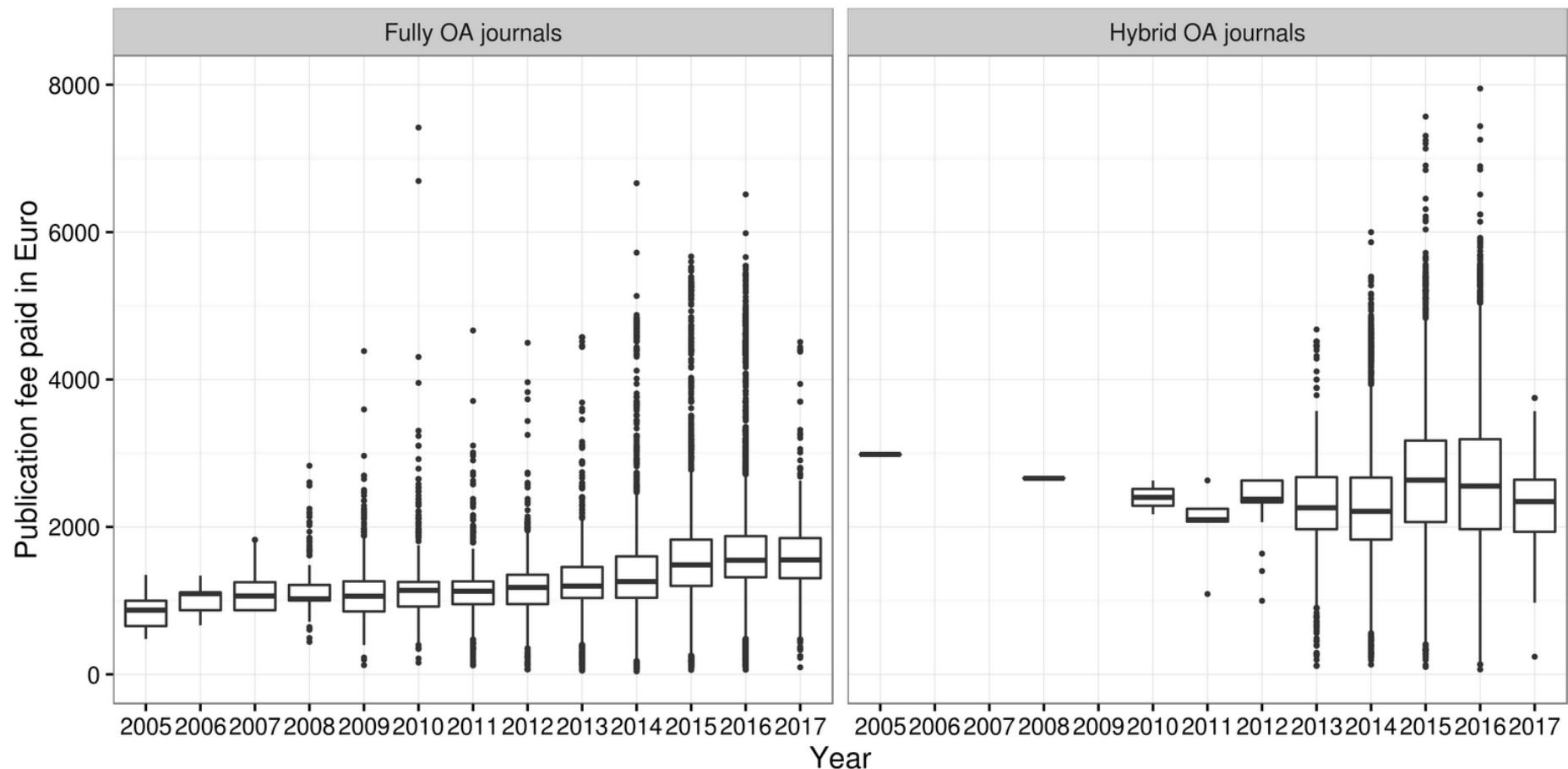
Downloads

Source code (zip)

Source code (tar.gz)

OpenAPC development (each October 2015-2017)

Year	Number of institutions	Number of articles	Amount in EURO	Median in EURO
2015 (10)	24	3.618	4.474.463	1.202
2016 (10)	40	18.370	32.802.770	1.544
2017 (10)	117	44.425	86.008.454	1.756
2018 (01)	150	46.011	88.212.005	1.724
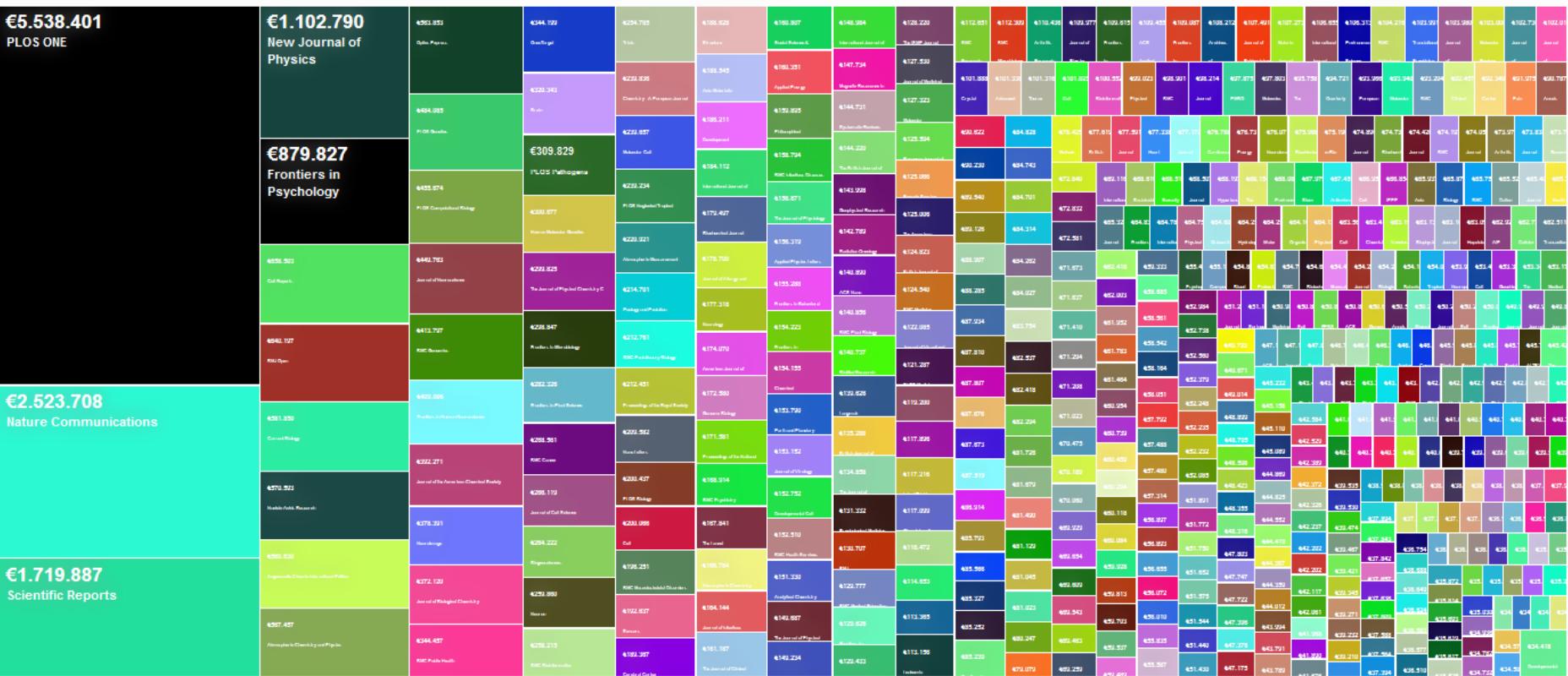
Spending distribution over fully and hybrid open access journals

Actual OpenAPC dataset

https://github.com/OpenAPC/openapc-de

Visualisation: https://treemaps.intact-project.org/apcdata/openapc/

€5.538.401 PLOS ONE
€1.102.790 New Journal of Physics
€879.827 Frontiers in Psychology
€2.523.708 Nature Communications
€1.719.887 Scientific Reports
€309.829 PLOS Pathogens

| Sort by: | Sum | Number of Articles | Mean Value | Standard Deviation | | Download as: | CSV | JSON |

Journals (4465 entries*)	Sum	Number of Articles	Mean Value	Standard Deviation	Percentage
🟩 PLOS ONE	€5.538.401	4438	€1.248	€234	6,44%
🟩 Nature Communications	€2.523.708	583	€4.329	€1.006	2,93%
🟩 Scientific Reports	€1.719.887	1244	€1.383	€428	2,00%
⬛ New Journal of Physics	€1.102.790	964	€1.144	€333	1,28%
🟪 Frontiers in Psychology	€879.827	619	€1.421	€495	1,02%
		➕ view small values			
Total	**€86.008.454**	**44425**	**€1.936**	**€979**	**100%**

Agenda

1. INTACT
2. OpenAPC approach
3. OpenAPC data analysis
4. **Open Access monitoring in Germany**

Open Access monitoring in Germany

Most comprehensive data set on APC payments, but far from being complete

Nearly all german universities, which take part in the DFG programme „Open Access publizieren", report data to OpenAPC (but data are not complete for the whole programme)

Comparing OpenAPC with publisher data shows, that research organisations and universities without institutional services have relevant APC expenditures as well

"Open Access in Deutschland" - Strategy of the Federal Ministry of Education and Research (BMBF), September 2016

Erfolgsmodelle von Open Access fördern

Das Bundesministerium für Bildung und Forschung wird die Entwicklung von Open Access-Initiativen und Projekten fördern und Erfolgsbeispiele mit einem Wettbewerb sichtbar machen.

Überzeugende innovative Ansätze, die eine weitere Verbreitung von Open Access an den Hochschulen und Forschungseinrichtungen ermöglichen, werden ausgezeichnet und ihre Umsetzung in den Wissenschaftsalltag gefördert. Im Fokus stehen hier Projekte, die die erfolgreiche Umsetzung von Open Access in die Praxis zeigen oder konkret die Publikation unter einem Open Access-Modell erleichtern.

Die leichte Auffindbarkeit von Open Access-Publikationen ist ein wichtiger Erfolgsfaktor für Open Access. Projekte wie die Bielefeld Academic Search Engine (BASE) zeigen, wie Open Access mit praktischen Projekten gefördert werden kann.

Vielfältige Projekte in der Wissenschaft, wie die Etablierung von Open Access-Zeitschriften, die Transformation von Subskriptionszeitschriften hin zu einem Open Access-Modell oder der Aufbau von Suchmaschinen und Datenbanken, welche die Nachnutzbarkeit, Auffindbarkeit oder vernetzte Verteilung von Open Access-Publikationen verbessern, zeigen, wie Open Access in der Praxis umgesetzt wird. Auch kleinere Projekte und Initiativen werden durch das Bundesministerium für Bildung und Forschung zukünftig unterstützt.

X. Open Access sichtbar und messbar machen – Open Access-Monitor

Das Bundesministerium für Bildung und Forschung wird einen Open Access-Monitor etablieren, der verlässlich den quantitativen Stand von Open Access in Deutschland verfolgen soll. Aufbauend auf einer Bestandsaufnahme der Open Access-Aktivitäten in Deutschland sollen weitere Handlungsbedarfe identifiziert und künftige Aktivitäten an diesen Erkenntnissen ausgerichtet werden. Zugleich können so die Effizienz künftiger Maßnahmen gemessen und Veränderungen des Publikationsverhaltens sichtbar gemacht werden.

Wenn Einrichtungen den Anteil ihrer Open Access-Veröffentlichungen beziffern können, so können sie auch die im Bereich Open Access weniger starken Bereiche identifizieren und Open Access zielgerichteter fördern.

Das Monitoring soll auch darstellen, aus welchen Quellen und in welcher Höhe Mittel für die wissenschaftliche Informationsversorgung und für die Finanzierung von Publikationen (sowohl Open Access als auch im subskriptionsbasierten Modell) aufgewandt werden. So lässt sich die Transformation zu Open Access zielgerichtet gestalten.

Open Access monitoring in Germany

BMBF „Förderrichtlinie des freien Informationsflusses in der Wissenschaft – Open Access", June 2017

➢ BMBF is funding 20 projects (press release 2017-12)
➢ Start of the first 16 projects
➢ FZ Jülich Zentralbibliothek: Synergien für Open Access - Open Access-Monitoring
➢ Funding of OpenAPC from BMBF will start 10/2018

Open Access monitoring in Germany

[Competence Center for Bibliometrics](#) (BMBF funded, connection to INTACT via I²SoS, Bielefeld U)

FZ Jülich Zentralbibliothek: [Open Access Barometer](#)

[DEAL](#) project: data acquisition, bibliometrics

National open access contact point oa2020-de

Bielefeld UL (BASE, OpenAPC)

Lessons learned from INTACT so far …

➢ Nothing is complete, need for trusted databases (DOAJ, Crossref, ISSN-L, Scopus, Web of Science, ….)

➢ Bibliometricians like terminated databases for calculations (basic totality of all academic articles cannot be exactly determined)

➢ Keep workflows as simple as possible for data providing institutions, enrich data with trusted identifiers

Lessons learned from INTACT so far …

➢ Check data quality not manually but automatically

➢ Measurement of open access shares should be contextualized (e.g. open access transformation, CRIS, funder policies, new deals and consortial structures)

➢ APC expenditures can be estimated also for those institutions, which don´t have central institutional APC services (WP for the next OpenAPC funding phase)

Lessons learned from INTACT so far …

➢ Use and demand for publisher reports to counter check data sources (Crossref, libraries, Web of Science, …)

➢ Tagging the open access status (license information) on the level of articles is unfortunately still a challenge for everyone involved in publication workflows

➢ Future potential of ORCID for pushing (open access) metadata into repositories?

Lessons learned from INTACT so far …

Data may be incomplete, but we have enough data to manage open access transformation and to develop a truly open access publication system

Thank you for your attention!

www.intact-project.org

Follow us: @oa_intact