Open-Access-Anteil bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin

Datenauswertung für das Jahr 2017

Andreas Hübner[1], Michaela Voigt[2], Pamela Finke[3], Christina Riesenweber[4]

Unter Mitarbeit von:

Eva Bunge (Deutsches Museum München)

Gabriele Börner (Charité - Universitätsmedizin Berlin)

Sebastian Dittmann (TU Berlin)

Sean Nowak (Freie Universität Berlin)

Bericht (Stand Dezember 2018): https://doi.org/10.14279/depositonce-7866

Daten (Stand September 2018): https://doi.org/10.14279/depositonce-7867

Dieses Material steht unter der Creative-Commons-Lizenz Namensnennung 4.0 International. Um den Text dieser Lizenz zu sehen, besuchen Sie https://creativecommons.org/licenses/by/4.0/.

[1] Open-Access-Büro Berlin, ORCiD: 0000-0001-7342-9789, andreas.huebner@open-access-berlin.de
[2] Technische Universität Berlin, ORCiD: 0000-0001-9486-3189, michaela.voigt@tu-berlin.de
[3] Humboldt-Universität zu Berlin, ORCiD: 0000-0001-9086-3202, pamela.finke@ub.hu-berlin.de
[4] Open-Access-Büro Berlin, ORCiD: 0000-0002-7449-9209, christina.riesenweber@open-access-berlin.de

Inhalt

Management Summary **1**

1 Einleitung **4**

2 Ergebnisse **5**

2.1 Artikel in genuinen Open-Access-Zeitschriften (Gold) 5

2.1.1 Berliner Korrespondenzautorschaft – Verteilung auf Institutionen und Verlage 6

2.1.2 Höhe der APCs 7

2.1.3 Verteilung der APCs auf Verlage 8

2.1.4 Artikel in genuinen Open-Access-Zeitschriften ohne APCs 9

2.2 Open-Access-Artikel in Hybridzeitschriften 9

2.2.1 Berliner Korrespondenzautorschaft – Verteilung auf Institutionen und Verlage 9

2.3 Open-Access-Artikel über den grünen Weg 10

2.3.1 Verteilung auf Repositorien 10

2.3.2 Verteilung auf Verlage (Ort der Erstpublikation) 11

3 Vorgehen **12**

3.1 Artikel in genuinen Open-Access-Zeitschriften (Gold) 13

3.2 Open-Access-Artikel in Hybridzeitschriften 14

3.3 Open-Access-Artikel über den grünen Weg 15

3.4 Datenaufbereitung 17

3.5 Grenzen der Methode 18

4 Anhang **21**

4.1 Übersicht Kennzahlen 21

4.2 Übersicht Datenbanken 22

4.3 Glossar 23

Management Summary

Die „Open-Access-Strategie für Berlin“ des Berliner Senats formuliert u. a. das Ziel, dass bis 2020 60 % aller Zeitschriftenartikel aus den öffentlichen Wissenschaftseinrichtungen des Landes Berlin im Sinne von Open Access zugänglich sein sollen.[5] Dazu ist ein Monitoring des gesamten Publikationsaufkommens wissenschaftlicher Zeitschriftenartikel sowie des entsprechenden Open-Access-Anteils nötig. Die vorliegende Studie liefert diese Daten für das Publikationsjahr 2017. Sie schließt damit an vorangegangene Studien an, die den Open-Access-Anteil der wissenschaftlichen Publikationen aus Berlin für die Jahre 2013–2015[6] und 2016[7] betrachtet haben. Mit diesen Studien kann Berlin als erstes deutsches Bundesland Zahlen für den landesweiten Open-Access-Anteil präsentieren.

Open-Access-Anteil in Berlin

In die Analyse einbezogen wurden die neun publikationsstärksten Wissenschaftseinrichtungen des Landes Berlin:

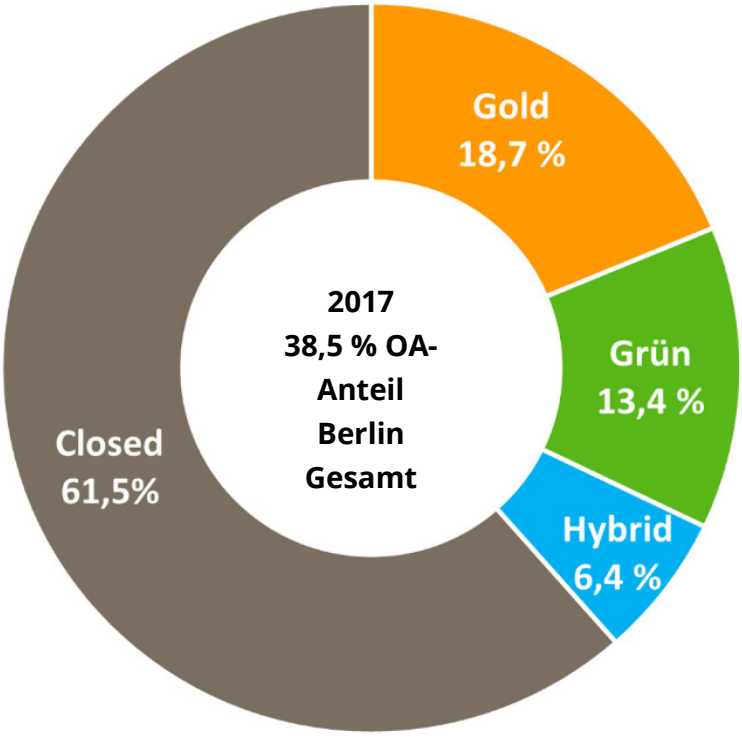
- Alice Salomon Hochschule (ASH)
- Beuth Hochschule (Beuth)
- Charité – Universitätsmedizin Berlin
- Freie Universität Berlin (FU)
- Hochschule für Technik und Wirtschaft Berlin (HTW)
- Hochschule für Wirtschaft und Recht Berlin (HWR)
- Humboldt-Universität zu Berlin (HU)
- Technische Universität Berlin (TU)
- Universität der Künste (UdK)

Der Gesamtanteil an Open-Access-Zeitschriftenartikeln mit Berliner Autorschaft im Publikationsjahr 2017 wird untergliedert in den Anteil von Artikeln in genuinen Open-Access-Zeitschriften („Gold Open Access“) und in Hybridzeitschriften („Hybrid Open Access“) sowie den Anteil von Open-Access-Zweitveröffentlichungen („Grün Open Access“).

---

[5] Vgl. http://www.open-access-berlin.de/strategie.
[6] Voigt, Michaela; Winterhalter, Christian (2016): Open-Access-Anteil bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin: Datenauswertung für die Jahre 2013–2015. https://doi.org/10.14279/depositonce-5570.
[7] Voigt, Michaela; Winterhalter, Christian; Riesenweber, Christina; Hübner, Andreas (2018): Open-Access-Anteil bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin: Datenauswertung für das Jahr 2016. https://doi.org/10.14279/depositonce-6866

Gesamtanzahl Zeitschriftenartikel:	10.923
Gesamtanteil Open Access:	4.206 (38,5 %)
Gold Open Access:	2.047 (18,7 %)
Grün Open Access:	1.464 (13,4 %)
Hybrid Open Access:	695 (6,4 %)



Gold Open Access

18,7 % der Zeitschriftenartikel im Publikationsjahr 2017 sind in genuinen Open-Access-Zeitschriften, also Open Access Gold, erschienen. Ca. die Hälfte dieser Artikel ist bei nur vier verschiedenen Verlagen erschienen, was auf eine Konzentration der Open-Access-Zeitschriften in wenigen Verlagen hindeutet. Mit der Verlagskonzentration geht zugleich eine Kostenkonzentration einher: Ca. 84 % der Article Processing Charges (APCs) wurden an insgesamt nur zehn Verlage gezahlt. Für gut ein Fünftel (17,8 %) der Artikel in Open-Access-Zeitschriften fielen keine APCs an.

Positiv zu bewerten ist, dass 81 % aller Open-Access-Gold-Artikel unter der freien Lizenz CC BY publiziert wurden. Damit wird nicht nur der freie Zugang, sondern auch die Nachnutzbarkeit der Forschungsergebnisse sichergestellt.

Grün Open Access

13,4 % der Zeitschriftenartikel im Publikationsjahr 2017 sind über den grünen Weg, also als Preprints oder Zweitpublikationen im Open Access auf Repositorien veröffentlicht worden. Ca. 62 % dieser Artikel wurden auf den Repositorien arXiv, Europe PubMed Central bzw. PubMed

Central abgelegt. Der hohe Anteil von arXiv ist als Indiz zu werten, dass Autor*innen vorzugsweise die in der jeweiligen Disziplin anerkannte und etablierte Publikationsinstanz nutzen. Der hohe Anteil auf (Europe) PubMed Central ist vermutlich darauf zurückzuführen, dass dieses Repositorium in zahlreichen Policies und Förderbestimmungen ausdrücklich als zur Zweitveröffentlichung vorgesehenes Repositorium benannt ist und dass es einen automatisierten Workflow zwischen Verlagen und (Europe) PubMed Central gibt.

Hybrid Open Access

Für das Publikationsjahr 2017 ließen sich 695 Open-Access-Artikel in Hybridzeitschriften identifizieren (6,4 % des ermittelten Gesamtpublikationsaufkommens), also Zeitschriften, die sich über Subskriptionen finanzieren, aber die Freischaltung einzelner Artikel durch Zahlung von APCs anbieten. Insgesamt verteilen sich die 695 Artikel auf 70 Verlage, wobei 57 % davon auf die drei Verlage Springer Nature, Elsevier und Wiley entfallen.

Die Methode

Diese Studie beruht auf der Auswertung von Daten aus 16 Literatur- und Zitationsdatenbanken. Sie beruht also nicht auf von den Wissenschaftseinrichtungen selbst erhobenen Daten, da eine solche einheitliche Erhebung im Moment nicht möglich ist. Die Bedingungen für die Datenerhebung, die methodischen Einschränkungen sowie deren Bedeutung für die Ergebnisse werden in dieser Studie beschrieben. So ist es möglich, die Ergebnisse dieser Studie mit anderen publizierten Open-Access-Quoten zu vergleichen bzw. die Vergleichbarkeit zu beurteilen.

Mittelfristig soll die Berliner Open-Access-Quote zusätzlich auf Grundlage von Daten aus den Berliner Einrichtungen selbst bestimmt werden, da dabei eine potenziell genauere Datenlage zur Verfügung stehen wird. Der Aufbau bzw. die Weiterentwicklung entsprechender Hochschulbibliographien an Berliner Einrichtungen ist derzeit in Arbeit.

Die in dieser Studie verwendete Methode ist publiziert und für andere Einrichtungen frei nachnutzbar (Dokumentation und Code).[8]

[8] https://github.com/tuub/oa-eval

1 Einleitung

Der Berliner Senat hat im Oktober 2015 eine „Open-Access-Strategie für Berlin“[9] verabschiedet, die unter anderem das Ziel formuliert, dass bis 2020 60 % aller Zeitschriftenartikel von Wissenschaftler*innen an Wissenschaftseinrichtungen des Landes Berlin im Sinne von Open Access zugänglich sein sollen.

Um geeignete Maßnahmen und Strategien zur Erhöhung des Open-Access-Anteils zu entwickeln und den voraussichtlichen Finanzierungsbedarf abzuschätzen, wurde 2016 eine Arbeitsgruppe aus Mitarbeiterinnen und Mitarbeitern der Bibliotheken von Charité, FU Berlin, HTW Berlin, HU Berlin und TU Berlin beauftragt, das Gesamtpublikationsaufkommen zu erfassen und den Open-Access-Anteil zu bestimmen. Der Bericht[10] und die zugrundeliegenden Daten[11] für den Publikationszeitraum 2013–2015 wurden Ende 2016 veröffentlicht. Die Anschlussstudie für das Publikationsjahr 2016[12] knüpft an die Methodik des Berichts für den Publikationszeitraum 2013–2015 an, wobei für die Detailanalyse methodische Verbesserungen erzielt werden konnten und erstmals zusätzlich der Publikationsanteil für Open Access Grün bestimmt wurde. Der vorliegende Bericht für das Publikationsjahr 2017 wendet im Wesentlichen die Methodik an, die auch für den Bericht für das Publikationsjahr 2016 verwendet wurde.

Für alle Studien gilt, dass ressourcenbedingt nur die neun publikationsstärksten Wissenschaftseinrichtungen des Landes Berlin in die Analyse des Publikationsaufkommens einbezogen wurden: Alice Salomon Hochschule, Beuth Hochschule, Charité – Universitätsmedizin Berlin, Freie Universität Berlin, Hochschule für Wirtschaft und Recht Berlin, Hochschule für Technik und Wirtschaft Berlin, Humboldt-Universität zu Berlin, Technische Universität Berlin und Universität der Künste. Zwei staatliche (Hochschule für Musik Hanns Eisler, Hochschule für Schauspielkunst Ernst Busch) und die privaten oder kirchlichen Berliner (Fach-)Hochschulen wurden nicht berücksichtigt.

In der Studie für die Publikationsjahre 2013–2015 fokussierte die Analyse auf die Bestimmung des „goldenen“ Open-Access-Anteils. Zusätzlich betrachtet wurden Open-Access-Artikel in Hybridzeitschriften, für die in einem experimentellen Verfahren Näherungswerte ermittelt wurden.

---

[9] https://www.parlament-berlin.de/ados/17/IIIPlen/vorgang/d17-2512.pdf

[10] Voigt, Michaela; Winterhalter, Christian (2016): Open-Access-Anteil bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin: Datenauswertung für die Jahre 2013–2015. https://doi.org/10.14279/depositonce-5570.

[11] Voigt, Michaela; Winterhalter, Christian (2016): Share of open access journal articles published by Berlin authors from 2013 to 2015: data. https://doi.org/10.14279/depositonce-5569.

[12] Voigt, Michaela; Winterhalter, Christian; Riesenweber, Christina; Hübner, Andreas (2018): Open-Access-Anteil bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin: Datenauswertung für das Jahr 2016. https://doi.org/10.14279/depositonce-6866

Die Ermittlung des „grünen“ Open-Access-Anteils stellte und stellt eine besondere Herausforderung dar.[13] Dank der Einführung des Webservices oaDOI[14] und der daraus entwickelten Anwendung „Unpaywall“ konnte für die vorliegende Studie der „grüne“ Open-Access-Anteil zumindest näherungsweise bestimmt werden.

2 Ergebnisse

Im Publikationsjahr 2017 wurden insgesamt 10.923 wissenschaftliche Artikel identifiziert, bei denen Angehörige der neun untersuchten Einrichtungen als Autor*innen auftreten. In 51,8 % der Fälle treten Berliner Wissenschaftler*innen als Erst- bzw. Korrespondenzautor*innen auf (im gesamten folgenden Text wird der Einfachheit halber verkürzt und nur noch von „Korrespondenzautor*innen“ gesprochen), für die verbleibenden Artikel sind sie Ko-Autor*innen. Im Anhang 4.1 werden die Open- bzw. Closed-Access-Anteile für alle Publikationen und zusätzlich unterschieden nach Zugehörigkeit der Korrespondenzautor*innen zu einer Berliner Einrichtung dargestellt.

2.1 Artikel in genuinen Open-Access-Zeitschriften (Gold)

2.047 Artikel der insgesamt 10.923 Artikel des Publikationsjahrs 2017 wurden in genuinen Open-Access-Zeitschriften veröffentlicht; dies entspricht einem Anteil von 18,7 %.

Diese Artikel sind bei 193 verschiedenen Verlagen erschienen. Es lässt sich eine deutliche Marktkonzentration erkennen: 73 % dieser Artikel sind bei zehn Verlagen erschienen, und ca. 50 % der Artikel sind bei nur vier verschiedenen Verlagen erschienen (Tab. 1). Diese Daten bestätigen Ergebnisse anderer Untersuchungen, die eine Konzentration der wichtigsten Open-Access-Zeitschriften auf wenige Verlage nachgewiesen haben.[15] Hinsichtlich der Marktkonzentration muss darüber hinaus – wie in *Kap. 3.4 Datenaufbereitung* dargestellt – berücksichtigt werden, dass

---

[13] Vgl. Voigt/Winterhalter (2016), S. 7: „Zum einen steht es Autorinnen und Autoren frei zu entscheiden, auf welchem Repositorium sie veröffentlichen – eine Analyse der jeweiligen institutionellen Repositorien ist daher nicht ausreichend. Zum anderen ist eine Zweitveröffentlichung aus rechtlichen Gründen häufig nur mit zeitlichem Verzug realisierbar; es stellt sich daher die Frage, ab wann ein Artikel als Open Access gezählt werden kann (ab de-facto OA-Verfügbarkeit oder ab Vorankündigung einer OA-Version für einen bestimmten Zeitpunkt). Eine andere Herausforderung ist die Datenlage insgesamt; insbesondere bei Preprints fehlt häufig die Verknüpfung zur späteren Verlagsveröffentlichung. Auch sind Titeländerungen nicht unüblich, was die Zuordnung der Closed-Access-Veröffentlichung zu einer (Preprint-)OA-Version erschwert.“ https://doi.org/10.14279/depositonce-5570.

[14] Vgl. http://unpaywall.org/faq. Der Dienst oaDOI wurde in der internationalen Open Access Week 2016 gestartet – kurz nach Abschluss der Untersuchung Open-Access-Anteil Berlin für die Jahre 2013–2015. Seit Januar 2018 wird dieser freie Webservice unter dem Namen „Unpaywall“ fortgeführt.

[15] Jahn, Najko; Tullney, Marco (2016): A study of institutional spending on open access publication fees in Germany. In: PeerJ 4, e2323. https://doi.org/10.7717/peerj.2323 (S. 10); Pinfield, Stephen; Salter, Jennifer; Bath, Peter A. (2016): The “total cost of publication” in a hybrid open-access environment. Institutional approaches to funding journal article-processing charges in combination with subscriptions. In: J Assn Inf Sci Tec 67 (7), S. 1751–1766. https://doi.org/10.1002/asi.23446 (S. 1759 f)

die als separate Imprints firmierenden Unternehmen Springer Nature und BioMed Central zu einer Verlagsgruppe gehören, deren Anteil somit insgesamt bei kumuliert 31,4 % der ermittelten Artikel in genuinen Open-Access-Zeitschriften liegt.

Tabelle 1 Verteilung von Gold-Open-Access-Artikeln (gesamt) auf die zehn Verlage mit den höchsten Publikationsanteilen.

Verlag	Reiner OA-Verlag	Anzahl Artikel	Anteil OA-Gold-Artikel
Springer Nature		412	20,1%
BioMed Central (BMC)	✓	232	11,3%
Frontiers Media S.A.	✓	226	11,0%
PLOS	✓	201	9,8%
MDPI	✓	160	7,8%
Elsevier BV		86	4,2%
Wiley		63	3,1%
Copernicus Publications	✓	46	2,2%
Oxford University Press (OUP)		37	1,8%
IOP Publishing		31	1,5%
Summe		*1.494*	*72,98 %*

2.1.1 Berliner Korrespondenzautorschaft – Verteilung auf Institutionen und Verlage

Wenn ein Artikel von mehreren Autor*innen verfasst wird, ist in der Regel ein*e Korrespondenzautor*in identifizierbar, der/die meist auch für die Finanzierung evtl. anfallender APCs zuständig ist. Für 1.020 der Berliner Artikel in genuinen Open-Access-Zeitschriften (also für 49,8 %) tritt ein*e Wissenschaftler*in der neun untersuchten Einrichtungen auch als Korrespondenzautor*in auf (Tab. 2).

Die Artikel mit Berliner Korrespondenzautorschaft verteilen sich auf 137 Verlage. Die zehn Verlage, bei denen die meisten Publikationen erschienen sind, decken 71,2 % des Gesamtpublikationsaufkommens ab (Tab. 3).

Tabelle 2 Affiliation der Berliner Korrespondenzautor*innen bei Gold-Open-Access-Artikeln.

Einrichtung	Anzahl Artikel	Anteil OA-Gold-Artikel
Charité	421	20,6%
FU	248	12,1%
TU	168	8,2%
HU	165	8,1%
FU; HU	8	0,4%
Beuth	3	0,1%
TU; HU	2	0,1%
TU; FU	1	<0,1%
TU; Charité	1	<0,1%
FU; UdK	1	<0,1%
HTW	1	<0,1%
ASH	1	<0,1%
Summe	*1.020*	*49,8%*

Tabelle 3 Verteilung der Gold-Open-Access-Artikel mit Berliner Korrespondenzautorschaft auf die zehn Verlage mit den höchsten Publikationsanteilen.

Verlag	Reiner OA-Verlag	Anzahl Artikel	Anteil OA-Gold-Artikel
Springer Nature		145	14,2%
Frontiers Media S.A.	✓	128	12,5%
PLOS	✓	116	11,4%
BioMed Central (BMC)	✓	108	10,6%
MDPI	✓	106	10,4%
Elsevier BV		36	3,5%
Wiley		29	2,8%
Copernicus Publications	✓	21	2,1%
Hindawi Limited	✓	20	2,0%
BMJ Publishing Group		17	1,7%
Summe		*726*	*71,2 %*

2.1.2 Höhe der APCs

Für ca. 82 % der Artikel in genuinen Open-Access-Zeitschriften (1.682 Artikel) wurden APCs an Verlage bezahlt. Da eine genaue Auflistung der bezahlten APCs durch die Einrichtungen nicht vorliegt, wurde die Höhe der vermutlich bezahlten APCs anhand der Angaben im DOAJ ermittelt. Die Zahlen stellen weder eine Übersicht über de-facto beglichene Gebühren dar, noch können sie als Basis für die Kalkulation eines validen Durchschnittswerts gelten. Sie können jedoch als Richtwert für einen Maximalbetrag gelten, den Berliner Einrichtungen im Jahr 2017 aus Haushalts- oder Drittmitteln für die Artikelgebühren in den jeweiligen Verlagen aufgebracht haben dürften.

Nicht berücksichtigt sind dabei Rabattierungsmöglichkeiten oder gestufte Gebühren für Art und Umfang der Artikel. Zur Frage allgemeiner Durchschnittswerte für APCs liegen neben dem ermittelten Wert aus den Vorjahresberichten zahlreiche Daten in OpenAPC[16] und Ergebnisse aus anderen Studien[17] vor.

Der durchschnittliche Wert für alle ermittelten APCs liegt bei 1.814,- € (1.682 Artikel, Umrechnungskurs vom 01.01.2017). Die geringste ermittelte APC beträgt 5,43 €, die höchste 4.752,95 €.

Die DFG hat als Förderhöchstgrenze in Publikationsfonds die Deckelung von 2.000,- € inkl. MwSt. festgelegt (1.680,- € zzgl. 19 % MwSt). Legt man Umrechnungskurse vom 01.01.2017 zugrunde, fallen 519 der ermittelten Artikel unter diese Grenze und wären somit prinzipiell förderfähig im Rahmen eines DFG-geförderten Open-Access-Publikationsfonds, mit einer durchschnittlichen APC in Höhe von 1.290,- € zzgl. 19 % MwSt. In Summe entspricht dies 669.444,95 € zzgl. 19 % MwSt.

2.1.3 Verteilung der APCs auf Verlage

Bei den zehn Verlagen mit den höchsten Publikationsanteilen von Gold-Open-Access-Artikeln (Tab. 1) sind APCs das zentrale (Open-Access-) Geschäftsmodell. Mit der Verlagskonzentration zugunsten des APC-Modells geht zugleich eine Kostenkonzentration einher: Ca. 84 % der Kosten sind bei diesen zehn Verlagen konzentriert. Bemerkenswert ist zudem, dass diese 84 % der Kosten für nur 73 % der Gold-Open-Access-Artikel angefallen sind. Tabelle 4 zeigt die Verteilung der ermittelten Beträge auf die zehn Verlage mit den höchsten ermittelten Beträgen.

Tabelle 4 Verteilung der ermittelten APCs auf die zehn Verlage mit den höchsten ermittelten Beträgen (=Summe APCs (exkl. 19 % MwSt.), Umrechnungskurs vom 01.01.2017).

Verlag	Anzahl Artikel	Summe APCs (€)	Durchschnittliche APC (€)
Springer Nature	412	638.922	1.551
Frontiers Media S.A.	226	594.955	2.633
BioMed Central (BMC)	232	388.505	1.675
PLOS	201	315.192	1.568
MDPI	160	202.806	1.268
Wiley	63	139.068	2.207
Elsevier BV	86	134.178	1.560
Oxford University Press (OUP)	37	64.615	1.746
BMJ Publishing Group	30	51.961	1.732
IOP Publishing	31	47.269	1.525

[16] https://treemaps.intact-project.org/

[17] Vgl. z.B. Jahn, Najko; Tullney, Marco (2016): A study of institutional spending on open access publication fees in Germany. In: PeerJ 4, e2323. https://doi.org/10.7717/peerj.2323; Pinfield, Stephen; Salter, Jennifer; Bath, Peter A. (2016): The "total cost of publication" in a hybrid open-access environment. Institutional approaches to funding journal article-processing charges in combination with subscriptions. In: J Assn Inf Sci Tec 67 (7), S. 1751–1766. https://doi.org/10.1002/asi.23446 sowie Solomon, David J.; Björk, Bo-Christer (2012): A study of open access journals using article processing charges. In: J Am Soc Inf Sci Tec 63 (8), S. 1485–1495. https://doi.org/10.1002/asi.22673

2.1.4 Artikel in genuinen Open-Access-Zeitschriften ohne APCs

Für knapp ein Fünftel der Artikel in genuinen Open-Access-Zeitschriften (365 Artikel, 17,8 %) fielen laut DOAJ keine APCs an. Betrachtet man nur die Open-Access-Artikel mit Berliner Korrespondenzautorschaft (179 Artikel), ist der Anteil sehr ähnlich (17,6 %). Es ist davon auszugehen, dass einige weitere, tendenziell bei kleineren Verlagen oder Institutionen erscheinende Open-Access-Zeitschriften ohne Publikationsgebühren existieren, die noch nicht im DOAJ gelistet sind und somit in dieser Untersuchung nicht erfasst werden können.

2.2 Open-Access-Artikel in Hybridzeitschriften

Für das Publikationsjahr 2017 ließen sich 695 Open-Access-Artikel in Hybridzeitschriften identifizieren (6,4 % des ermittelten Gesamtpublikationsaufkommens), also Zeitschriften, die sich über Subskriptionen finanzieren, aber die Freischaltung einzelner Artikel durch Zahlung von APCs anbieten. Aufgrund der für die Publikationsjahre 2016 und 2017 weiterentwickelten Erhebungsmethodik sind die Zahlen nur sehr eingeschränkt mit den eher experimentell ermittelten Zahlen für die Jahre 2013–2015 vergleichbar. Insgesamt verteilen sich die 695 Artikel auf 70 Verlage, wobei 57 % davon auf die drei Verlage Springer Nature, Elsevier und Wiley entfallen (Tab. 5).

Tabelle 5 Verteilung der Hybrid-Open-Access-Artikel (gesamt) auf die zehn Verlage mit den höchsten Publikationsanteilen.

Verlag	Anzahl Artikel	Anteil Artikel
Springer Nature	166	23,9%
Elsevier BV	159	22,9%
Wiley	73	10,5%
Royal Society of Chemistry (RSC)	32	4,6%
Impact Journals, LLC	31	4,5%
BMJ Publishing Group	26	3,7%
American Physical Society (APS)	24	3,5%
IOP Publishing	20	2,9%
American Chemical Society (ACS)	13	1,9%
Oxford University Press (OUP)	13	1,9%
Summe	*557*	*80,1%*

2.2.1 Berliner Korrespondenzautorschaft – Verteilung auf Institutionen und Verlage

Für 278 der 695 Open-Access-Artikel in Hybridzeitschriften konnte eine Berliner Korrespondenzautorschaft ermittelt werden (40 %; Tab. 6). Insgesamt verteilen sich diese 278 Artikel auf 46 Verlage, wobei 54 % auf die drei Verlage Elsevier, Springer Nature und Royal Society of Chemistry entfallen (Tab. 7).

Tabelle 6 Affiliation der Berliner Korrespondenzautor*innen bei Hybrid-Open-Access-Artikeln.

Einrichtung	Anzahl Artikel	Anteil Artikel
Charité	101	14,5%
FU	70	10,1%
TU	68	9,8%
HU	31	4,5%
HTW	3	0,4%
UdK	2	0,3%
FU; HU	1	0,1%
HWR	1	0,1%
ASH	1	0,1%
Summe	*278*	*40,0%*

Tabelle 7 Verteilung der Hybrid-Open-Access-Artikel mit Berliner Korrespondenzautorschaft auf die zehn Verlage mit den höchsten Publikationsanteilen.

Verlag	Anzahl Artikel	Anteil Artikel
Elsevier BV	70	25,2%
Springer Nature	54	19,4%
Royal Society of Chemistry (RSC)	25	9,0%
Wiley	24	8,6%
Impact Journals, LLC	11	4,0%
BMJ Publishing Group	10	3,6%
American Society for Microbiology	7	2,5%
American Physical Society (APS)	5	1,8%
Oxford University Press (OUP)	5	1,8%
American Chemical Society (ACS)	4	1,4%
Summe	*215*	*77,3%*

2.3 Open-Access-Artikel über den grünen Weg

Für das Publikationsjahr 2017 betrug der Anteil von Open-Access-Artikeln über den grünen Weg, also als Open-Access-Preprint oder -Zweitveröffentlichung, am Gesamtpublikationsaufkommen zum Zeitpunkt der Datenauswertung 13,4 % (1.464 Artikel).

2.3.1 Verteilung auf Repositorien

Die Open-Access-Artikel, die als Grün-Open-Access-Artikelversionen vorliegen, sind auf 237 verschiedenen Repositorien verteilt, wobei mitunter verschiedene Versionen eines Artikels auf einem oder mehreren Repositorien verfügbar sind. Tabelle 8 zeigt die zehn Repositorien, über die am häufigsten Grün-Open-Access-Artikelversionen verfügbar sind. Dabei lässt sich eine starke

Konzentration beobachten: Ca. 62 % der Grün-Open-Access-Artikel sind auf arXiv, Europe PubMed Central und/oder PubMed Central abgelegt (Tab. 9).

Tabelle 8 Rangfolge der zehn Repositorien, über die am häufigsten Grün-Open-Access-Artikelversionen verfügbar sind (Mehrfachnennungen möglich, da es mitunter parallele Versionen auf einem oder mehreren Repositorien gibt).

Repositorium	URL	Anzahl Versionen	Anteil Versionen
arXiv	arxiv.org	628	25,84%
Europe PMC[18]	europepmc.org	344	14,16%
PubMed Central	www.ncbi.nlm.nih.gov	120	4,94%
EconStor	www.econstor.eu	85	3,50%
University College of London Discovery	discovery.ucl.ac.uk	56	2,30%
MPG.PuRe	pubman.mpdl.mpg.de	45	1,85%
HAL	hal.archives-ouvertes.fr	44	1,81%
Academic Bibliography Ghent University	biblio.ugent.be	40	1,65%
King's College London Research Portal	kclpure.kcl.ac.uk	37	1,52%
DESY Publication Database	bib-pubdb1.desy.de	37	1,52%

Tabelle 9 Anzahl Grün-Open-Access-Artikel, die über die Top-3-Repositorien verfügbar sind. Insgesamt handelt es sich um 901 verschiedene Artikel. Diese Artikel sind in einigen Fällen in mehreren der drei genannten Repositorien vorhanden.

Repositorium	Anzahl Artikel	Anteil Artikel
arXiv (arxiv.org)	567	38,73%
Europe PubMed Central (europepmc.org)	336	22,95%
PubMed Central (www.ncbi.nlm.nih.gov)	114	7,79%
Anzahl Grün-Open-Access-Artikel		
arXiv und/oder Europe PubMed Central und/oder PubMed Central	901	61,54%

2.3.2 Verteilung auf Verlage (Ort der Erstpublikation)

Tabelle 10 zeigt die Rangfolge der zehn Verlage, bei denen die Berliner Grün-Open-Access-Artikel als Erstpublikation am häufigsten erschienen sind.

Gemessen am Publikationsaufkommen von Berliner Wissenschaftler*innen bei den drei großen kommerziellen Verlagen Elsevier, Springer Nature und Wiley sind die relativen Anteile der „grünen“ Versionen von subskriptionspflichtigen Artikeln aus diesen Verlagen mit 10–12 % (je nach Verlag) gering.

[18] Partner von PubMed Central, vgl. https://www.ncbi.nlm.nih.gov/pmc/about/pmci/

Tabelle 10 Rangfolge der zehn Verlage, bei denen die Berliner Grün-Open-Access-Artikel als Erstpublikation am häufigsten erschienen sind.

Verlag	Anzahl Artikel	Anteil Artikel
Elsevier BV	226	15,4%
Springer Nature	221	15,1%
American Physical Society (APS)	153	10,5%
Wiley	126	8,6%
Oxford University Press (OUP)	57	3,9%
Informa UK Limited	56	3,8%
IOP Publishing	50	3,4%
American Chemical Society (ACS)	48	3,3%
SAGE Publications	43	2,9%
AIP Publishing	35	2,4%
Summe	*1.015*	*69,3%*

3 Vorgehen

Für die Analyse des Publikationsaufkommens des Jahres 2017 der untersuchten Berliner Einrichtungen wurden bibliographische Daten zu den Dokumenttypen „Article“ und „Review“[19] aus 16 Literatur- und Zitationsdatenbanken ermittelt und exportiert (vgl. *4.2 Übersicht berücksichtigte Datenbanken*).

Zur Ermittlung der Zugehörigkeit zu einer Berliner Institution wurde in den Datenbanken, die eine Affiliationssuche ermöglichen, nach verschiedenen Namensvarianten der beteiligten Institutionen gesucht.

Die Abfragen in den Datenbanken erfolgten im Zeitraum 2.–30.8.2018. Die bibliographischen Daten wurden mithilfe eines Python-Skripts automatisiert ausgewertet; dessen Funktionsweise wird in *Kap. 3.1 Artikel in genuinen Open-Access-Zeitschriften (Gold)* näher erläutert. Das Skript steht auf GitHub zur Nachnutzung zur Verfügung.[20] Zur Verifizierung der Angaben zu Open-Access-Artikeln in hybriden Zeitschriften wurden zusätzlich mithilfe von OpenRefine[21] die Schnittstellen von Crossref und Unpaywall abgefragt (vgl. *Kap. 3.2 Open-Access-Artikel in Hybridzeitschriften*).

Die bibliographischen Daten aus den Datenbanken wurden in das oben erwähnte Skript geladen und aggregiert, normalisiert und auf Dubletten geprüft.

[19] Berücksichtigt wurden – analog zu vergleichbaren Untersuchungen – die Publikationstypen „(Research) Article“ und „Review (Article)“, also Artikel, die den aktuellen Forschungsstand in einem Fachgebiet zusammenfassen. Der Publikationstyp „Book review“, d. h. Rezension, ist hingegen nicht berücksichtigt.
[20] Python-Skript für die Analyse des Open-Access-Anteils vgl. https://github.com/tuub/oa-eval, hier verfügbar ist auch eine detaillierte (technische) Dokumentation des Skripts.
[21] OpenRefine ist eine freie Software zur Datentransformation und -aggregation, s. http://openrefine.org/.

3.1 Artikel in genuinen Open-Access-Zeitschriften (Gold)

Als Open Access Gold bezeichnen wir Zeitschriftenartikel, die in genuinen Open-Access-Zeitschriften erscheinen, also am Ort und zum Zeitpunkt der Erstveröffentlichung frei verfügbar sind. Bei Open Access Gold unterscheidet man grundsätzlich zwischen Finanzierung der Zeitschrift mit Artikelgebühren (Article Processing Charges, APCs) und Finanzierungsmodellen ohne APCs.

Open Access Gold mit APCs

Viele Open-Access-Zeitschriften weltweit stellen den Autor*innen bzw. deren Einrichtungen Publikationsgebühren in Rechnung. Ihre Höhe variiert je nach Verlag, Zeitschrift und Fachgebiet. Im Rahmen einer institutionellen Mitgliedschaft bei Open-Access-Verlagen zahlen Forschungsinstitutionen eine Jahresgebühr, um den Institutionsangehörigen die Veröffentlichung ihrer Forschungsergebnisse in einer Open-Access-Zeitschrift über Sammelrechnungen oder zu rabattierten APCs zu ermöglichen.

In Berlin gibt es derzeit an vier Einrichtungen zentrale Fonds zur Übernahme von APCs, die unter Eigenbeteiligung der Hochschulen durch das DFG-Förderprogramm „Open Access Publizieren“[22] finanziert werden (FU seit 2012, TU seit 2017, Charité und HU seit 2018).

Open Access Gold ohne APCs

Zahlreiche Open-Access-Zeitschriften finanzieren sich nicht über APCs, sondern über institutionelle Grundförderung, Finanzierung durch Fachgesellschaften oder Drittmittel. Viele Publikationsangebote von wissenschaftlichen Einrichtungen, z. B. Universitätsverlage, operieren nach diesem Modell.

Auf die Aggregation der Artikeldaten folgte die Identifikation von Artikeln in genuinen Open-Access-Zeitschriften. Zentrales Nachweismittel für Open-Access-Zeitschriften ist das Directory of Open Access Journals (DOAJ). Das DOAJ stellt seine Daten als Datei[23] und über eine Webschnittstelle[24] zur Verfügung. Angaben zu ISSNs in dieser Liste wurden mit den aggregierten Artikeldaten abgeglichen. Für Artikel, für die in den Datenbanken eine DOI, jedoch keine ISSN nachgewiesen wird, wurden ISSNs über die DOI-Registrierungsagentur Crossref ermittelt und mit DOAJ-Daten abgeglichen.

Für einen Großteil der Open-Access-Artikel fallen Publikationsgebühren (APCs) an. Da APCs in der Regel von den Korrespondenzautor*innen („corresponding authors“) getragen werden, wurden in einem dritten Schritt diejenigen Artikel identifiziert, für die Angehörige der Berliner Hochschulen Erst- oder Korrespondenzautor*innen[25] sind. Einige Datenbanken liefern keine bzw.

---

[22] DFG-Förderprogramm „Open Access Publizieren“, vgl. http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/open_access_publizieren/
[23] Download der DOAJ-Metadaten: https://doaj.org/csv
[24] Webschnittstelle des DOAJ: https://doaj.org/api/v1/docs
[25] Die Rolle von Erst- bzw. Korrespondenzautor*innen unterscheidet sich je nach Disziplin – teils signalisiert die Stellung (erste oder letzte Stelle) in der Autor*innenliste, wer primär Ansprechpartner*in für einen Artikel ist, teils wird bei der Autorschaft alphabetisch gelistet und die Korrespondenzautoschaft separat

uneinheitlich strukturierte Angaben hierzu. Daher erstellt das Skript eine Liste von Artikeln, deren Korrespondenzautorschaft manuell zu überprüfen ist.[26]

Angaben zur Höhe der APCs für Artikel in Open-Access-Zeitschriften basieren auf Angaben im DOAJ (Stand: 5.9.2018); sie wurden nicht manuell verifiziert. Die Umrechnung in Euro erfolgte auf Basis der Wechselkurse unter XE.com.[27] Zu beachten ist, dass die in diesem Bericht aufgeführten Preisangaben Nettowerte exklusive 19 % Mehrwertsteuer sind.

3.2 Open-Access-Artikel in Hybridzeitschriften

„Hybrid" werden subskriptionsbasierte Zeitschriften genannt, die einzelne Artikel gegen Bezahlung Open Access verfügbar machen. Hier werden also das Modell der kostenpflichtigen Subskription und das Modell der Zahlung von APCs miteinander kombiniert, so dass für Forschungseinrichtungen doppelte Kosten entstehen. Aus diesem Grund ist z. B. im Rahmen einer DFG-Förderung die Erstattung von Gebühren für Hybrid Open Access aus Publikationsfonds nicht erlaubt. Da Hybridzeitschriften selbst nicht genuin Open Access sind, ist die Auffindbarkeit der hybriden Open-Access-Artikel häufig erschwert.[28]

Zur Untersuchung von Open-Access-Artikeln in Hybridzeitschriften[29] wurde ausgehend von den bibliographischen Daten aus den Datenbanken für alle verfügbaren DOIs eine Anfrage an die Schnittstelle des Webservices Unpaywall[30] gestellt. Zusätzlich wurde die Crossref-Schnittstelle[31] abgefragt und die Metadaten überprüft. Als „hybrid" wird ein Artikel dann gewertet, wenn es laut Unpaywall bzw. Crossref eine Open-Access-Version gibt, die über den Verlag, und zwar mit einer

---

ausgewiesen. Zum Teil weisen Datenbanken die Korrespondenzautorschaft aus, so dass die Angaben vom Skript automatisiert verarbeitet werden können. Fehlt die Angabe zur Korrespondenzautorschaft, werden im genutzten Skript Affiliationsangaben der Erstautorenschaft ausgewertet. Eine Analyse von Eva Bunge zeigte, dass Erst- und Korrespondenzautorschaft in zwei Dritteln der Fälle einhergehen, so dass es zu einer Fehlerrate von ca. 5 % kommt, wenn in Ermangelung einer Angabe zur Korrespondenzautorschaft die Affiliation der Erstautorschaft ausgewertet wird. Vgl. Bunge, Eva. (2017). Ermittlung von Open-Access-Kennzahlen – Ein Bericht aus der Praxis. Open-Access-Tage 2017. https://doi.org/10.5281/zenodo.999482 (S. 9).

[26] Dass Erst- oder Korrespondenzautor*innen für die Organisation der Finanzierung verantwortlich sind, ist eine notwendige Grundannahme für diese und andere Auswertungen. Für verbessertes Monitoring und die damit verbundene Kostenkalkulation besteht Optimierungsbedarf für Datenerfassung und -auswertung (geteilte Korrespondenzautorschaften, unspezifische Funding-Angaben etc.), etwa durch die Einführung eines Feldes „OA paying affiliation", siehe Gumpenberger C, Hölbling L and Gorraiz JI (2018): On the Issues of a „Corresponding Author" Field-Based Monitoring Approach for Gold Open Access Publications and Derivative Cost Calculations. Front. Res. Metr. Anal. 3:1. https://doi.org/10.3389/frma.2018.00001.

[27] Die Umrechnung in Euro erfolgte auf Basis der Wechselkurse vom 1.1.2017 (Annäherung an de-facto-gezahlte APC-Kosten) bzw. 22.09.2018 (Annäherung an anfallende APC-Kosten nach aktuellem Kursstand) unter http://www.xe.com.

[28] Inzwischen wurden in der Open-Access-Community Tools entwickelt, die das Identifizieren von OA-Artikeln in Hybridzeitschriften erleichtern: https://najkoja.shinyapps.io/hybridoa/ (Basis: OpenAPC-Daten und Crossref-Metadaten) bzw. https://github.com/ryregier/OAcounts (Basis: Crossref-Metadaten sowie Angaben zum OpenAccess-Status von Unpaywall).

[29] Der Einfachheit halber wird nachfolgend von Hybridzeitschriften gesprochen, wenn es sich um Subskriptionszeitschriften handelt, die Autor*innen gegen Gebühr die OpenAccess-Veröffentlichung einzelner Artikel ermöglichen.

[30] Webschnittstelle von Unpaywall: https://unpaywall.org/api/v2

[31] Webschnittstelle von Crossref: http://api.crossref.org/

Creative-Commons-Lizenz versehen, zugänglich ist und es sich nach Abgleich mit dem DOAJ nicht um eine genuine OA-Zeitschrift handelt.

Für jeden „Treffer“ wurde zudem die ISSN der Zeitschrift gespeichert und durch Abfrage der DOAJ-Schnittstelle geprüft, ob es sich bei der Zeitschrift nicht doch um eine genuine Open-Access-Zeitschrift handelt. Dieser Fall kann eintreten, wenn der Artikel wegen fehlender oder fehlerhafter ISSN beim ersten Abgleich mit den DOAJ-Daten nicht als Open-Access-Artikel erkannt wurde.

Dieses Vorgehen entspricht dem des Vorjahres, so dass ein direkter Vergleich zu 2016 möglich ist. Allerdings ist ein Vergleich mit den ermittelten Zahlen für die Jahre 2013–2015 nicht zu empfehlen, da die Erhebungsmethodik für die Publikationsjahre 2016 und 2017 abweicht.

3.3 Open-Access-Artikel über den grünen Weg

Als den Grünen Weg des Open Access verstehen wir Open Access über die zusätzliche Open-Access-Veröffentlichung eines Artikels, der zunächst in einer Subskriptionszeitschrift publiziert wurde (Zweitveröffentlichung oder Preprint). Die parallele Veröffentlichung erfolgt auf einem Repositorium. Da es auf dem maschinengestützten Weg nicht gesichert möglich ist, zwischen der Preprint- und der Postprint-Version eines Artikels in einem Repositorium zu unterscheiden, werden in dieser Untersuchung auch Preprints als Open Access Grün gezählt, wenn, wie oben beschrieben, eine in einer Subskriptionszeitschrift publizierte Version vorliegt.[32]

Um die Anzahl der Artikel zu ermitteln, die über den grünen Weg Open Access verfügbar sind, wurden Daten von Unpaywall ausgewertet: Als „grün“ wurden die Artikel gewertet, für die eine Open-Access-Version über ein Repositorium verfügbar ist. Dabei wurden auch Preprints auf Repositorien berücksichtigt, d. h., wenn das ursprünglich eingereichte Manuskript, bevor es einen Begutachtungsprozess durchlaufen hat, Open Access verfügbar ist.

In der vorliegenden Studie wurden für den Anteil der „grünen“ Artikel nur solche Artikel berücksichtigt, die über ein gesichertes Repositorium verfügbar sind. Nicht berücksichtigt wurden Versionen in akademischen Netzwerken wie ResearchGate oder Academia.edu. Diese Angebote werden von Autor*innen aufgrund ihrer Funktionen zur akademischen Vernetzung zwar gerne genutzt. Mitunter werden jedoch nicht die Verlagsbedingungen für Zweitveröffentlichungen eingehalten, was etwa zu Klagen von Elsevier und ACS gegen ResearchGate führte – mit dem Resultat, dass der Zugriff auf zahlreiche Volltexte auf ResearchGate gesperrt wurde.[33] Andere

---

[32] Wir folgen damit der Definition der Unpaywall-Betreiber: „Green OA: Green articles are published in a toll-access journal, but self-archived in an OA archive. These 'OA archives' are either disciplinary repositories like ArXiv, or institutional repositories (IRs) operated by universities, and the archived articles may be either the published versions, or electronic preprints (Harnad et al., 2008). Most Green OA articles do not meet the BOAI definition of OA since they do not extend reuse rights (making them Gratis OA).“ s. Piwowar H, Priem J, Larivière V, Alperin JP, Matthias L, Norlander B, Farley A, West J, Haustein S. (2018) The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles. PeerJ 6:e4375 https://doi.org/10.7717/peerj.4375

[33] vgl. Dalmeet Singh Chawla (6.10.2017): Publishers take ResearchGate to court, alleging massive copyright infringement http://www.sciencemag.org/news/2017/10/publishers-take-researchgate-court-alleging-massive-copyright-infringement

Studien berücksichtigen bei der Messung des Open-Access-Anteils auch solche Versionen[34]; aufgrund der Instabilität dieses Zugriffs haben wir uns gegen dieses Vorgehen entschieden.

Auf Repositorien abgelegte Versionen von Artikeln in genuinen Open-Access-Zeitschriften oder Versionen von Open-Access-Artikeln in Hybridzeitschriften wurden nicht berücksichtigt, um Doppelzählungen zu vermeiden.

Die Auswertung der zehn Repositorien mit den jeweils größten Anteilen an Berliner OA-Grün-Artikeln (Tab. 7) erfolgte auf Basis der in Unpaywall nachgewiesenen Links zu einzelnen Artikelversionen, wobei der Link auf die Domain gekürzt wurde. Für 66 Versionen auf Repositorien (Achtung: nicht 66 Artikel) wird in Unpaywall ein Handle-Link mit globalem Handle-Resolver als Domain (hdl.handle.net) nachgewiesen, so dass keine eindeutige Zuordnung der Repositorien-URL bei der Datenauswertung möglich ist. Diese 66 Versionen wurden bei der Auflistung der Top-Repositorien daher nicht berücksichtigt.

Der Open-Access-Anteil über den grünen Weg ist das dynamischste Segment am Publikationsaufkommen, die Ermittlung dieses Anteils muss stets mit einem exakten Zeitstempel versehen sein. Aufgrund der verzögerten Zugänglichmachung der Artikel in Repositorien durch die Autor*innen und aufgrund geltender Embargo-Regelungen von durchschnittlich zehn Monaten[35] wird dieser Anteil sukzessive ansteigen und möglicherweise in Zukunft den größten Anteil am Open-Access-Publikationsaufkommen ausmachen. Der Anteil „grün" kann jedoch nie final ermittelt werden: Idealerweise steigt er kontinuierlich an, wenn Autor*innen ihre Publikationen sukzessive zweitveröffentlichen. Insofern ist der angegebene Open-Access-Anteil über den grünen Weg als Mindestwert zu verstehen.

Zusätzlich ist davon auszugehen, dass aufgrund zunehmender „Open-Access-Grün"-Forderungen der Forschungsförderer, der Entwicklung von Open-Access-Tools (z. B. dissemin[36]), Projekten wie

---

[34] So etwa eine Studie zum Open-Access-Anteil in Großbritannien, vgl. Monitoring the transition to open access: December 2017: http://www.universitiesuk.ac.uk/policy-and-analysis/reports/Documents/2017/monitoring-transition-open-access-2017.pdf

[35] Bezogen auf OpenAccess-Publikationen im Kontext von Horizon 2020 liegt eine Analyse zur durchschnittlichen Embargo-Dauer vor: „The average embargo period of green OA publications is 10 months, that is a decrease of 1 month from the 2016 sample. 40 % of articles have an embargo period of 11–12 months, followed by 575 articles (or 33 % with no embargo period at all. 302 articles, that is 17 % have an embargo period of 12,1-24 months and 162 articles or 9 % of 0,1 to 6 months. Finally, 12 articles, that is 1 %, have an embargo period that is longer than 36 months."
https://data.europa.eu/euodp/data/dataset/open-access-to-scientific-publications-horizon2020

[36] Ein ebenfalls vergleichsweise neuer Dienst, der bibliographische Quellen und Informationen aus SHERPA-RoMEO (http://www.sherpa.ac.uk/romeo/) auf einer ansprechend gestalteten und intuitiv nutzbaren Oberfläche zusammenführt, ist die in Frankreich entwickelte Plattform dissemin (https://dissem.in). Über diesen Service lässt sich die für eigene Publikationen prüfen, welche davon bereits frei verfügbar sind (die freie Verfügbarkeit wird hier sehr weit ausgelegt und z. B. auch ResearchGate einbezogen) bzw. für welche der noch nicht frei verfügbaren Publikationen die Möglichkeit zur Zweitveröffentlichung besteht. Besonders umfassend funktioniert diese Prüfung, sofern man sich mit einem bestehenden ORCID-Account anmeldet, da in der Folge alle dort hinterlegten Publikationen in dissemin abgeprüft werden können. Über ein persönliches Profil besteht eine bequeme Möglichkeit zum Upload von Publikationen u. a. nach HAL (https://hal.archives-ouvertes.fr) oder Zenodo (https://zenodo.org).

DeepGreen[37] oder OA-EZB[38] und die Verbesserung von Zweitveröffentlichungs-Workflows in den Institutionen[39] die Anzahl „grüner“ Artikelversionen höher und die Zeitspanne bis zu ihrer Verfügbarkeit geringer wird.

3.4 Datenaufbereitung

DOIs, bei denen die Abfragen der Schnittstellen von Unpaywall und Crossref Fehler gemeldet haben, wurden geprüft und ggf. korrigiert. Für die korrigierten DOIs wurden die Schnittstellen von Unpaywall und Crossref erneut abgefragt und die Ergebnisse in der Auswertung berücksichtigt.

Schreibweisen von beispielsweise Verlagen und Journalen sind in den bibliographischen Datenbanken und auch in den Crossref-Daten oft heterogen. Um die Daten auswerten zu können, wurde die Schreibweise von Zeitschriftentiteln und Verlagen vereinheitlicht. Sofern möglich wurden Cluster nach Verlagsgruppe gebildet: Einige Verlagsgruppen publizieren unter verschiedenen Imprints (Bsp. Wiley: „Wiley and Sons“, „Wiley-Blackwell“, „Wiley-VCH“) bzw. sind Dienstleister für Fachgesellschaften (Bsp. Wiley: AGU, IUCr). Eine Auswertung nach Verlagen sollte jedoch nach Verlagsgruppen erfolgen, um tatsächliche Konzentrationen am Markt identifizieren zu können. Es wurden insgesamt vier Verlagscluster identifiziert: Wiley, Springer Nature, Wolters Kluwer und IOP Publishing. Für diese Verlagscluster ist eine Besonderheit zu beachten: Zwar gehört BioMed Central (BMC) zur Verlagsgruppe Springer Nature. Bei BMC erscheinen – anders als bei anderen Springer-Imprints – ausschließlich genuine Open-Access-Zeitschriften; zudem bietet BMC eigene Prepaid- bzw. Mitgliedschaftsmodelle an. Aufgrund dieser besonderen Stellung wird BMC bei der Clusterung separat ausgewiesen.

Auch für die Daten zur Korrespondenzautorschaft wurden Bereinigungen vorgenommen: Zum einen wurden Fälle, in denen das Skript eine geteilte Berliner Korrespondenzautorschaft (bspw. „TU Berlin“ und „Beuth“) lieferte, manuell verifiziert und bei Bedarf korrigiert. Zum anderen wurden die in den Daten enthaltenen E-Mail-Adressen manuell ausgewertet: Dazu wurde nach Artikeln gefiltert, für die keine Berliner Korrespondenzautorschaft identifiziert wurden, nach E-Mail-Domains der untersuchten Einrichtungen gesucht (z. B. „fu-berlin.de“), die Korrespondenzautorschaft auf Basis des Primärtextes geprüft und ggf. der Eintrag für die Berliner Korrespondenzautorschaft ergänzt.

Um Angaben zu Repositorien granularer auswerten zu können, wurden die Unpaywall-Daten zudem besonders aufbereitet: Aus dem Unpaywall-Feld „oa_location“ wurden alle Angaben zu Repositorien extrahiert und pro Artikel in einem separaten Feld abgelegt. Für die Auswertung

---

[37] „Das Projekt DeepGreen will wissenschaftliche Veröffentlichungen, die lizenzrechtlich nach Embargofristen frei zugänglich gemacht werden dürfen, in den Open Access überführen. Hierbei konzentriert sich das Projekt [...] auf speziell von der Deutschen Forschungsgemeinschaft (DFG) geförderte, überregional verhandelte Lizenzen, die sog. Allianz-Lizenzen, welche spezielle Open-Access-Regelungen enthalten.“ vgl. https://deepgreen.kobv.de

[38] „Ziel des von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projektes „OA-EZB: Open-Access-Services der Elektronischen Zeitschriftenbibliothek“ ist daher der Ausbau und nachhaltige Betrieb von vernetzten und aggregierten Mehrwertdiensten, um die Akzeptanz von Open Access sowohl bei den Endnutzern als auch bei den Autoren und Autorinnen signifikant zu erhöhen.“ vgl. https://www.uni-regensburg.de/bibliothek/projekte/oa-ezb/index.html

[39] Blasetti, A. et al. (2018): Smash the Paywalls: Workflows und Werkzeuge für den grünen Weg des Open Access. Poster: https://doi.org/10.5281/zenodo.1469080, Preprint für Artikel in Informationspraxis: http://hdl.handle.net/10419/184668

nach Versionen auf verschiedenen Repositorien wurde dieses Feld in Sub-Felder aufgeteilt. In der zusätzlichen Datendatei werden so einzelne Open-Access-Versionen des jeweiligen Artikels separat aufgeführt, so dass nunmehr eine Auswertung danach möglich ist,

a) auf wie vielen verschiedenen Repositorien ein Artikel im Durchschnitt parallel verfügbar ist,
b) wie häufig Unpaywall eine Open-Access-Version einem Berliner Repositorium zuordnen kann oder
c) wie häufig bestimmte Fachrepositorien oder Preprint-Server genutzt werden.

Es ist also zu beachten, dass die Gesamtzahl der hier aufgeführten Open-Access-Versionen nicht der Gesamtzahl der Artikel entspricht.

3.5 Grenzen der Methode

Datenbasis

Grundsätzlich sind die Ergebnisse dieser Studie abhängig von der genutzten Datenbasis. Artikel in Zeitschriften, die nicht in einer der für diese Untersuchung genutzten geprüften Datenbanken indexiert sind (siehe Anhang 4.2 Übersicht berücksichtigte Datenbanken), werden nicht berücksichtigt.

Affiliation

Voraussetzung für die Identifizierung von Artikeln der Berliner Wissenschaftler*innen ist das Erfassen der institutionellen Zugehörigkeit (Affiliation) in den betrachteten externen Datenbanken. Häufig wird in den Datenbanken nur eine Affiliation erfasst; gehört ein*e Autor*in mehreren Institutionen an, bleibt ein Artikel ggf. unentdeckt. Zahlreiche (deutsche) Institutionen haben inzwischen Affiliationsrichtlinien verabschiedet. Von den in dieser Studie betrachteten Einrichtungen haben bisher die FU Berlin, die TU Berlin und die Charité eine solche Richtlinie verabschiedet; an der HU Berlin sind entsprechende Planungen weit vorangeschritten. Zudem hat die Hochschulrektorenkonferenz im April 2018 „Leitlinien zur Nennung von Affiliationen bei Publikationen“ veröffentlicht.[40] Daneben könnten internationale Entwicklungen im Bereich von Verfahren für die Personenidentifikation[41] und Organisationenidentifikation[42] mittel- bis langfristig zur Vereinfachung der Methodik bzw. zur Reduktion der Fehlerquote beitragen.

---

[40] Hochschulrektorenkonferenz: Leitlinien zur Nennung von Affiliationen bei Publikationen – Empfehlung der 24. HRK-Mitgliederversammlung, https://www.hrk.de/positionen/beschluss/detail/leitlinien-zur-nennung-von-affiliationen-bei-publikationen/

[41] Am Markt haben sich mehrere Systeme zur eindeutigen Autorenidentifzierung etabliert: ORCiD (Betreiber: gemeinnützige Organisation; Verwendung nicht beschränkt auf bestimmte Plattform), ResearcherID (Betreiber: Clarivate Analytics, Anwendung in Web of Science) sowie ScopusID (Betreiber: Elsevier; Verwendung in Scopus), wobei ORCiD aufgrund der Offenheit des Systems ein besondere Rolle zukommt.

[42] Etabliert sind bislang Ringgold (vgl. https://www.ringgold.com/ringgold-identifier/) und GRID (Global Research Identifier Database, vgl. https://www.grid.ac/). Unter der Federführung von ORCiD, DataCite und Crossref hat sich eine Arbeitsgruppe gegründet, die analog zu ORCiD für die Identizifierung von Personen auch ein offenes System zur Identifizierung von Organisationen aufbauen und dabei auch auf die beiden genannten Identifier Ringgold und GRID zurückgreifen möchte (vgl. https://orcid.org/content/organization-identifier-working-group).

Der Aspekt der verlässlichen Identifizierung der Affiliation wirkt sich auf Ergebnisse in internationalen Hochschulrankings aus: Vergleiche wie etwa das Times Higher Education University Ranking[43] greifen für die Auswertung von Publikationsleistungen (als ein Aspekt des Leistungsvergleichs) auf Daten aus Web of Science und Scopus zurück. Fehlen hier (bzw. in den Publikationen selbst) passende Affiliationsangaben oder ist die Abfolge oder Zuordnung bei multiplen Affiliationen einer Autorin bzw. eines Autors nicht korrekt, werden die Publikationen nicht berücksichtigt, so dass die Institutionen im Gesamtergebnis schlechter abschneiden.

Dublettenerkennung

Bei der Dublettenerkennung können Fehler auftreten, da im Web of Science alle Titel, unabhängig von der Originalsprache des Artikels, ins Englische übersetzt werden. Ist ein Titel in einer anderen Datenbank mit dem Titel der Originalsprache erfasst und wird für diesen Artikel keine DOI nachgewiesen, greift die Dublettenerkennung im Skript nicht. Die Publikation würde doppelt gezählt.

Eine potentielle Fehlerquelle für den Dublettenabgleich ist zudem die Länge der Autor*innennamen: Ist ein Name sehr kurz bzw. enthält er nur wenige Konsonanten, kann es zu Fehlern kommen. Im Rahmen der Datenbereinigung wurde nach identischen Titeln gesucht und Angaben wurden manuell verglichen. Dennoch kann nicht vollständig ausgeschlossen werden, dass Dubletten enthalten sind, da es zeichenweise Abweichungen bei den Titeln in den Datenbanken geben kann.

Identifikation von Open-Access-Zeitschriften

Der Identifikation von Open-Access-Zeitschriften sind Grenzen gesetzt, wenn sie (noch) nicht im DOAJ erfasst sind oder wenn in den bibliographischen Daten der Datenbanken die ISSN fehlt. Das führt dazu, dass Artikel in diesen nicht entdeckten Open-Access-Zeitschriften auch nicht als Open-Access-Artikel berücksichtigt werden können. Es ist bekannt, dass einige Zeitschriften die Artikel frei verfügbar machen, aber aus formalen oder inhaltlichen Kriterien nicht im DOAJ gelistet werden.

Währungsschwankungen

Wie in Kap. 2.1.2 Höhe der Kosten von APCs gezeigt, sind Währungsschwankungen ein maßgeblicher Einflussfaktor für APC-Kosten. Dies verdeutlicht auch die Relevanz dieses in der Studie für die Publikationsjahre 2013–2015 angeführten Parameters bei der Kostenkalkulation und Formulierung von Kriterien für die Vergabe von Fördermittel (z. B. hochschulinterne Publikationsfonds).

Open Access Grün

Es ist aktuell nicht ermittelbar, welche Version (d. h. Preprint, akzeptiertes Manuskript oder Verlagsversion) durch welchen Autor bzw. welche Autorin zweitveröffentlicht wurde: Rückschlüsse sind aufgrund der URL zum Dienst im Fall von institutionellen Repositorien möglich, aber insbesondere Open-Access-Versionen, die über Fachrepositorien oder über institutionelle Angebote verfügbar sind, lassen keinen Rückschluss zu, durch welche* Autor*in die Zweitveröffentlichung

[43] vgl. https://www.timeshighereducation.com/world-university-rankings/2018/world-ranking

vorgenommen wurde. Außerdem kann momentan nicht ermittelt werden, wann die Zweitveröffentlichung erfolgt ist.

Open Access in verschiedenen Disziplinen

Verzichtet wurde im Bericht auf eine Analyse der disziplinspezifischen Verteilung der Open-Access-Publikationen, deren methodischen Probleme und Grenzen bereits im Bericht für die Publikationsjahre 2013–2015 aufgezeigt worden waren. Auswertungen hierzu müssen daher weiterhin als Desiderat gelten, wobei jedoch zugleich auf neuere Ergebnisse aus übergreifenden Erhebungen zurückgegriffen werden kann.[44]

Die Datensituation für die sozial- und vor allem geisteswissenschaftlichen Fachgebiete ist jedoch weiterhin unbefriedigend – in vielen einschlägigen Datenbanken fehlt nach wie vor die Möglichkeit einer affiliationsbezogenen Suche. Dies lässt sich mittel- und langfristig nur durch eine systematische Erfassung aller Publikationen nach vereinbarten Standards an den Hochschulen selbst lösen (z. B. gemeinsame Berliner Hochschulbibliographie sowie abgestimmte Affiliationsrichtlinien), wie dies auch als eines der Ziele in der Berliner OA-Strategie formuliert ist.

[44] Science-Metrix (2018): Analytical Support for Bibliometrics Indicators: Open access availability of scientific publications. http://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix_open_access_availability_scientific_publications_report.pdf sowie Martín-Martín, A., Costas, R., van Leeuwen, T. & López-Cózar, E. D. (2018). Evidence of Open Access of scientific publications in Google Scholar: a large-scale analysis. https://doi.org/10.17605/OSF.IO/FSUJY

4 Anhang

4.1 Übersicht Kennzahlen

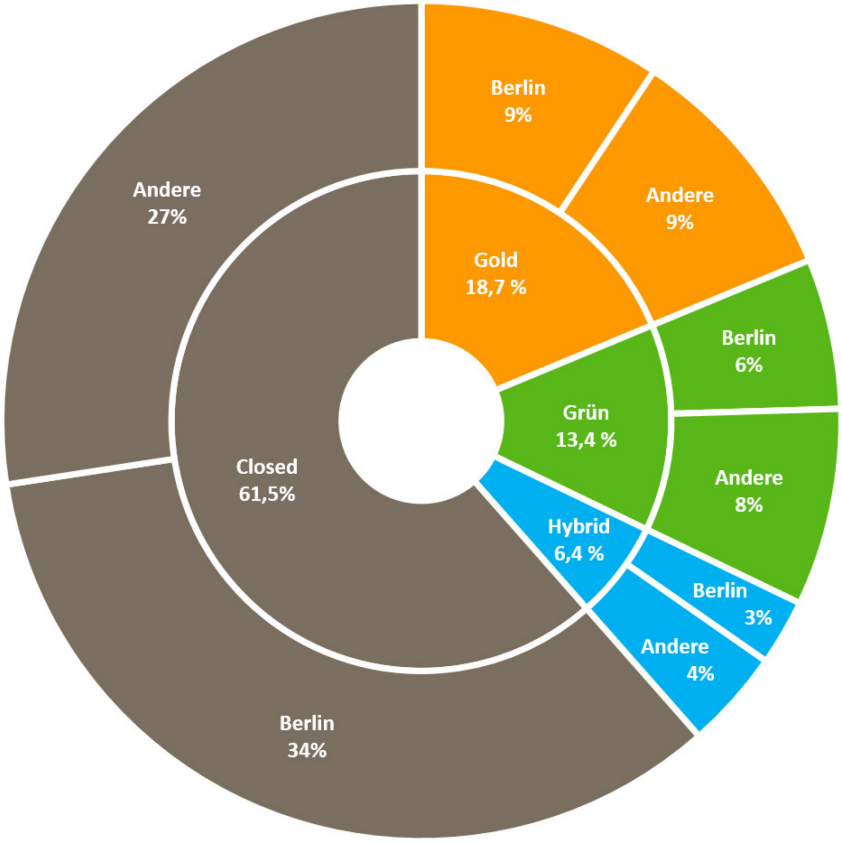


Abbildung A1 Übersicht von Open-Access- und Closed-Access-Anteilen bei Zeitschriftenartikeln von Wissenschaftlerinnen und Wissenschaftlern an Einrichtungen des Landes Berlin sowie Angabe der Anteile Berliner Erst- und Korrespondenzautorschaften.

Tabelle A1 Übersicht der Open-Access-Kennzahlen.

	Korrespondenz-autorschaft	Anzahl Artikel	Anzahl Artikel in %	Verteilung Berlin/Andere in %
Gold-Open-Access	Berlin	1.020	9,34%	49,8%
	Andere	1.027	9,40%	50,2%
	Gesamt	*2.047*	*18,74%*	
Grün-Open-Access	Berlin	633	5,80%	43,2%
	Andere	831	7,61%	56,8%
	Gesamt	*1.464*	*13,40%*	
Hybrid-Open-Access	Berlin	278	2,55%	40,0%
	Andere	417	3,82%	60,0%
	Gesamt	*695*	*6,36%*	
Closed	Berlin	3.722	34,07%	55,4%
	Andere	2.995	27,42%	44,6%
	Gesamt	*6.717*	*61,49%*	
Summe		*10.923*	*100,00%*	

4.2 Übersicht Datenbanken

Folgende Datenbanken wurden bei der Analyse berücksichtigt:

- Web of Science Core Collection
- Inspec
- SciFinder (CAPlus)
- PubMed
- TEMA
- IEEE Xplore
- ProQuest Social Sciences
- Business Source Complete
- GeoRef
- CAB Abstracts
- CINAHL
- Academic Search Ultimate
- Embase
- Library and Information Science Abstracts (LISA)
- Scopus
- SPORTDiscus

4.3 Glossar

Article Processing Charges (APC), häufig auch Publikationsgebühren, Page Charges oder Publication Charges genannt, sind Gebühren, die bei Annahme eines Zeitschriftenbeitrags zur Publikation anfallen. Sie werden üblicherweise von der beschäftigenden Einrichtung stellvertretend für die Autorin bzw. den Autor bezahlt.

Ein **Digital Object Identifier** (DOI) ist eine eindeutige Identifikationsnummer für digitale Objekte, mithilfe derer eine dauerhafte Verlinkung möglich ist. Eine vergleichbare technische Lösung ist das URN-System. Das DOI-System ist international sehr bekannt und wird für die sog. persistente Adressierung am häufigsten genutzt.

Goldener Weg oder **Open Access Gold** meint die unmittelbare und freie Zugänglichkeit zu einer Veröffentlichung am originären Publikationsort. Bei diesem Publikationsort kann es sich beispielsweise um eine wissenschaftliche Zeitschrift, einen Konferenzband, einen Sammelband oder eine Monografie handeln.

Grüner Weg oder **Grün Open Access** bezeichnet die Veröffentlichung von qualitätsgesicherten Textpublikationen und anderen digitalen Objekten auf einem institutionellen oder disziplinären Repositorium. Dazu zählen im Sinne dieses Berichts Preprints oder Zweitveröffentlichungen von Verlagsversionen einer Veröffentlichung bzw. akzeptierten Manuskripten.

Eine **Hybridzeitschrift** ist eine subskriptionsbasierte Zeitschrift, die einzelne Artikel gegen Bezahlung Open Access verfügbar machen. Eine solche Option wird von Verlagen unter Begriffen wie „Author's Choice" oder „OpenChoice" beworben.

Die **Korrespondenzautorin** bzw. der **Korrespondenzautor** („corresponding author") ist Hauptansprechpartner*in für einen Artikel gegenüber Verlag und Leserschaft. Mitunter wird die Korrespondenzautorschaft nicht gesondert ausgewiesen, oftmals übt dann die Erstautorin bzw. der Erstautor diese Funktion aus. Open-Access-Publikationsgebühren werden in der Regel von den Erst- oder Korrespondenzautoren (bzw. deren Arbeitgebern) getragen.

Eine **Open-Access-Zeitschrift** ist eine Zeitschrift, bei der alle Artikel sofort am originären Publikationsort frei zugänglich sind. In dem vorliegenden Bericht sind damit die Zeitschriften gemeint, die auch im Directory of Open Access Journals (DOAJ) gelistet sind und damit bestimmten formalen Qualitätsansprüchen (z. B. in Bezug auf Zugänglichkeit oder wissenschaftlichen Begutachtungsverfahren) genügen.

Der Begriff **Preprint** wird in der vorliegenden Studie definiert als ursprüngliche Manuskriptversion der Autorin bzw. des Autors, d. h. die Version, welche bei einer Zeitschrift eingereicht wird. Zu unterscheiden davon sind die akzeptierte Manuskriptversion (Version, in die Änderungen aus der wissenschaftlichen Begutachtung eingeflossen sind) und die Verlagsversion. Mitunter werden unter Preprint auch solche publizierten Beiträge gefasst, die nie bei einer Zeitschrift eingereicht werden oder nicht zur Veröffentlichung angenommen werden. In der vorliegenden Studie werden jedoch nur Artikel berücksichtigt, die auch in wissenschaftlichen Zeitschriften publiziert wurden und somit in Literatur- und Zitationsdatenbanken indexiert werden.

Bei **Publikationsfonds** handelt es sich um Mittel, die von wissenschaftlichen Einrichtungen verwaltet werden und die dazu dienen, Article Processing Charges zu finanzieren.

Ein **Repositorium** ist ein mit dem Internet verbundener Server zur geordneten Bereithaltung und Archivierung elektronischer Daten. Dabei kann es sich um wissenschaftliche Publikationen von Texten oder Forschungsdaten sowie um Daten des kulturellen Erbes handeln.

Subskriptionsmodell: Finanzierung von Zeitschriften oder regelmäßig erscheinenden Publikationen durch Abonnements bzw. Zugriffslizenzen. Häufig sind nicht mehr einzelne Zeitschriften, sondern ganze Zeitschriftensegmente eines Verlagsprogramms Gegenstand einer Subskription (so genannte Bundle Deals).