Gray OA 2012-2016:
Open Access Journals Beyond *DOAJ*

Walt Crawford

A Special Issue of

Cites & Insights

Crawford at Large/Online Edition

Libraries • Policy • Technology • Media

Volume 17, Number 1: January 2017 ISSN 1534-0937

Contents

1. The Big Picture

> How many open access (OA) articles are published each year? How many open access (OA) journals publish how many OA articles? What proportion of those journals and articles involve fees (usually called Article Processing Charges or APCs)? How much did each article cost?

That's the first paragraph of *Gold Open Access Journals 2011-2015* (henceforth *GOAJ*), which went on to answer those questions *for serious gold open access*, where "serious" was defined by inclusion in the *Directory of Open Access Journals* (henceforth *DOAJ*). But there's more to OA, even to gold OA.

Comprehensive answers to those questions may not be feasible, for a variety of reasons, but this report should get a lot closer to the full picture—by adding "gray OA": gold OA journals that are *not* in *DOAJ*. (This does not include journals dropped from *DOAJ* in mid-2016: those were covered in the earlier report.)

Herewith, then, some oversimplified figures for gray OA, offered comparably to those on page 1 of *GOAJ:*

- Gray OA journals published 155,347 articles in the first half of 2016. Full-year figures include 296,963 articles in 2015; 255,183 in 2014; 188,645 in 2013; and 125,039 in 2012. Extrapolating for 2016, this shows decreasing rate of growth in each year.
- In all, 7,743 gray OA journals published at least one article between January 1, 2012 and June 30, 2016, so you *could* say there were an average of 20 articles per journal in 2015—but that's misleading.
- There are a staggering 18,910 journal *titles* in the gray OA world as defined for this report—but most of those titles were never anything more than titles and template-generated webpages.
- Among the 6,841 journal clearly stating charges, 6,374 (93%) *do* charge author-side charges (henceforth APC), and those journals published 96.7% of the 2015 articles—up slightly from 96.5% in 2014. (There are another 902 journals that apparently charge but

don't state the charges; including those journals, no-APC journals accounted for only 2.9% of 2015 articles and 3.1% of 2014 articles, with APC-charging journals accounting for 97.1$ and 96.9% respectively.) In short, nearly all of gray OA involves APCs.

- Including only journals that actually published articles in a given year or half-year, and excluding failed-to-state journals, 4,631 journals published 139,327 articles in the first half of 2016, taking in a *maximum* of $46,418,625 in APCs—or $333 per article ($343 excluding articles in no-fee journals). That's an average of 30 articles per journal, which extrapolates to 60 for the full year. For 2015, the comparable numbers are 5,252 journals with articles; 262,398 articles; $81,130,347 maximum APCs; $309 average cost per article ($320 excluding no-fee journals); and 50 articles per journal.

These numbers are all far too simple because they treat gray OA as a homogeneous whole, which is not at all the case. This report explores the leftover portion of OA on some detail and looks at some issues with a previous report based on sampling the gray OA universe. As appropriate, I'll include *GOAJ* figures and grand totals for 2012-2015—noting that such totals still aren't quite comprehensive. Still: nearly 864,000 articles in gold OA journals in 2015: that's a striking number, more than a third of the presumed 2.5 million total scholarly articles per year.

The Gray OA Universe

How did I unearth the gray OA universe? I tried to make brandy out of sour grapes: I used two lists compiled by a librarian whose antipathy toward all open access, and eagerness to label and shame any and all "ppppredatory" journals and publishers, make the lists essentially useless for their intended purpose but quite promising for the purposes of this report.

That's right: this report is based on publishers and journals in Jeffrey Beall's lists as of July 8, 2016. Here's the process:

1. I copied the two lists on July 8, 2016. Using a saved copy of *DOAJ* from January 1, 2016, I trimmed a very small number of matching journal titles from the list of "independent" journals, leaving 902 journals.
2. After eliminating exact duplicates from the publisher list, I attempted to visit each of the remaining 1,025 "publishers."

Where publisher sites existed and weren't malware (or aliases for other publisher sites), I prepared a set of journal titles and URLs. I also checked for publisher-level APCs and for stated country, noting each if present.

3. I then visited (or attempted to visit) each journal website, determine the APC (if that wasn't provided at the publisher level), and count the articles for 2012, 2013, 2014, 2015, and the first six months of 2016 (a sometimes approximate figure). One exception: for some of the "publishers" I think of as "template publishers," each with 390 or more "journals" with essentially identical webpages, I sorted the journals in alphabetic order, then checked the first 100. If I found no articles at all in the first hundred journals, I assumed all "journals" were empty—so I may have missed a few tiny journals. (Since no template publisher checked in its entirety, as most were, had more than 111 articles *total* in 2014 or more than 36 articles in any other year, I'm reasonably confident there's no serious undercount.)
4. In the process, I verified at least once per publisher or "independent" journal that I could in fact open an article as a PDF or full-text HTML view and noted apparent cases of papermills—journals with apparently random subject coverage and typically improbably short review turnaround times, usually with spiking article counts in one or two years.
5. During this process, I checked journals against *DOAJ*, removing (and not rechecking) a total of 527 journals. Those journals are not part of this report.

This was a long process. Many of these publishers and "publishers" have sites that are difficult to deal with, and some sites appeared and disappeared. If numbers sometimes don't quite add up between portions of the report, that's probably why. When I've cross-checked, discrepancies have always been trivial (e.g., less than 0.5%)

In general, this is a quantitative report, not a qualitative one, as is discussed more in Chapters 2-4.

Journals and publishers were omitted for various reasons, discussed further in Chapter 2. It's fair to note that only 547 publishers had one or more non-*DOAJ* journals with at least one published article in the last 4.5 years.

There are so few non-APC journals in the gray world that I'm not doing free-vs-pay tables and graphs in most cases, substituting gray-vs-DOAJ in some cases.

The Biggest Numbers

Note that, unlike *GOAJ*, I *do* include journals with hidden/unstated APCs in some discussions because they represent a larger portion of the whole: for 2015, 12.2% of the active journals and 11.6% of the articles, compared to 1.1% and 1.5% for *GOAJ*. I'll use the abbreviation UA for these Unknown APC journals and note where they are and aren't included—for example, they're *not* included in discussions of revenue, subjects or countries.

	Journals	**Active 2015**	**Articles**	**Art/Jrnl**
Free	467	384	8,712	22.7
Pay	6,374	4,868	253,686	52.1
UA	902	736	34,565	47.0
Gray Total	**7,743**	**5,988**	**296,963**	**49.6**
GOAJ	10,324	9,531	566,922	59.5
OA Total	18,067	15,519	863,885	55.7

Table 1.1. Journals and articles, overall

Table 1.1 shows the key figures for gray journals for 2015 (the last full year counted) and, for comparison, the serious OA figures as reported in *GOAJ*. Note that some journals don't publish articles every year (a *lot* of gray OA journals don't!) and that gray OA journals were on average smaller than serious OA journals…while the handful of free gray OA journals were generally *very* small.

Table 1.2 shows article counts for journals counted in this report, with codes for a number of special cases.

Code	Count	Jan-Jn 2016	2015	2014	2013	2012
A	3,765	137,291	246,053	207,704	147,519	96,216
B3	350				2,164	1,695
B4	549			2,826	2,491	1,367
B5	1,205		10,646	8,143	5,665	3,964
BC	104	2	502	402	514	661
BF	844	1,230	4,239	3,971	3,830	2,771
BR	24	804	958	917	749	538
UA	902	16,020	34,565	31,220	25,713	17,827
Total	7,743	155,347	296,963	255,183	188,645	125,039

Table 1.2. Articles per year and codes

Notes on the codes (other than "UA," already explained):

- "**A**" is the catchall code for journals that didn't get any other code.
- "B3" journals haven't published any articles since 2013, and can probably be considered defunct.
- "B4" journals haven't published any articles since 2014. They might be failing or on hiatus.
- "B5" journals published articles in 2015 but not in the first half of 2016. Some of these have very long lead times for posting articles. (Most counts were taken in August, September and October 2016.)
- "BC" journals fall into one of two categories: explicitly ceased or merged into other journals (thus the 2013-2015 numbers), or with no articles more recent than 2012. It seems fair to assume that a journal with no activity in 3.5 years is defunct.
- "BF" journals have either one or two 2016 articles, too few to represent robust publishing.
- "BR" journals consist entirely or primarily of reviewed papers presented at conferences.

These codes are directly comparable to those used in *GOAJ* (where "UA" was coded "CA"). There are proportionally more of most "B" codes; gray OA journals are more erratic in general.

Journal Stability

Let's look at the erratic nature of gray OA—and, for comparison, add some new data for *GOAJ*. To wit: how many journals manage to publish a significant number of articles for more than one year? How many do so for three or more years?

That raises the question of what's a significant number of articles—and I've seen answers as high as 40, which seems extreme. For this discussion, we'll use two figures: more than four (or, for the first half of 2016, more than two) and, later, more than nine (with no special provision for 2016).

Years > 4	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	334	563	10.7%	532	466	189
One	678	666	23.4%	491	345	166
Two	857	1,068	43.7%	566	375	173
Three	757	874	60.4%	889	434	176
Four	750	826	76.1%	827	828	267
Five	1,255	1,255		1,255	1,255	1,255
Total	4,631	5,252		4,560	3,703	2,226

Table 1.3a. Gray journals publishing five or more articles per year

Table 1.3a shows the number of gray OA journals (coded A or B, not including UA) that actually published articles in each year, broken down by the number of years a journal published at least five articles (or at least three for January-June 2016).

If you define two active years as minimal for a stable journal, most gray journals make it: more than three out of four. But if four years is the target, only 40% manage. Note also that, out of 6,841 A&B journals, there's never a year without at least 1,589 not publishing *any* articles.

Table 1.3b shows articles in those journals—and makes the data much more interesting. (This table also explains the decimal point in Table 1.3a percentages: it's there because one percentage in Table 1.3b rounds to zero.)

To wit: the 23% of journals with no more than one good year published only 2% of the 2015 articles—and even lower percentages in 2014 and 2013. The quarter of journals that were around since 2012

and always published at least five articles a year accounted for 55% of all articles in 2015.

Years > 4	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	440	1,113	0.4%	1,007	898	415
One	5,646	4,603	2.2%	2,837	2,254	1,620
Two	15,797	18,848	9.4%	5,629	4,520	2,362
Three	23,073	42,015	25.4%	24,968	5,168	2,471
Four	25,250	51,604	45.1%	45,167	22,998	5,694
Five	69,121	144,184		144,355	127,094	94,650
Total	139,245	262,367		223,963	162,932	107,212

Table 1.3b. Articles in gray OA journals publishing more than four each year

How do these figures compare to similar measures for *GOAJ*?

Years > 4	2015	Cum%	2014	2013	2012	2011
None	38	0.4%	46	46	43	42
One	243	2.9%	206	155	117	106
Two	660	9.9%	722	331	231	181
Three	985	20.2%	1,067	1,153	531	374
Four	1,189	32.7%	1,467	1,467	1,475	761
Five	6,416		6,416	6,416	6,416	6,416
Total	9,531		9,924	9,568	8,813	7,880

Table 1.3c. *GOAJ* journals with more than four articles per year

There's no "special count for 2016" here—and using 2015 for comparison, we see that more than 90% of the journals published five or more articles in at least three years, and that more than two-thirds did so in all five years: *GOAJ* journals are *much* more stable than gray journals. (Note also that, in 2014, only 400 or 4% of ongoing journals didn't publish any articles, compared to 23% for gray OA's best year.) Only 3% were what I think of as one-shot wonders, journals with only one good year (or new journals in 2015), compared to more than 23% of gray journals.

Table 1.3d, directly comparable to Table 1.3b, shows an even more dramatic difference: journals stable for all five years account for nearly eight out of ten articles in 2015 and higher percentages in earlier years.

Those stable for four years or more account for more than 88% of all 2015 articles (compared to 40% for gray OA).

Years > 4	2015	Cum%	2014	2013	2012	2011
None	95	0.02%	132	124	112	123
One	4,096	0.74%	1,105	857	524	1,260
Two	24,293	5.02%	17,868	2,961	2,650	1,986
Three	37,143	11.58%	34,477	25,250	7,184	4,969
Four	52,581	20.85%	57,917	53,639	39,356	14,119
Five	448,714		448,537	410,644	388,818	337,892
Total	566,922		560,036	493,475	438,644	360,349

Table 1.3d. Articles in GOAJ journals publishing at least five articles per year

For the sake of completeness, Tables 1.3e and 1.3f show the same information as Tables 1.3a and 1.3b, but for the 903 UA journals, those with hidden or absent APCs.

Years > 4	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	44	96	13.0%	51	43	41
One	91	94	25.8%	61	43	20
Two	89	139	44.7%	99	51	34
Three	105	135	63.0%	143	70	35
Four	106	125	80.0%	126	126	51
Five	147	147		147	147	147
Total	582	736		627	480	328

Table 1.3e. Gray UA journals with at least five articles per year

Although the numbers are much smaller, the patterns are similar. That's also true for Table 1.3f.

Years > 4	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	59	209	0.6%	123	94	86
One	435	646	2.5%	353	223	123
Two	994	2,144	8.7%	883	567	325
Three	2,542	6,580	27.7%	4,812	1,083	603
Four	4,140	8,035	51.0%	6,584	7,091	1,858
Five	7,850	16,951		18,465	16,655	14,832
Total	16,020	34,565		31,220	25,713	17,827

Table 1.3f. Articles in UA gray journals with five or more articles per year

Raising the Bar

Years > 9	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	1,209	1,512	28.8%	1,238	952	429
One	739	870	45.4%	630	463	212
Two	635	747	59.6%	559	400	200
Three	629	674	72.4%	683	438	222
Four	603	633	84.5%	634	634	347
Five	816	816		816	816	816

Table 1.4a. Gray journals publishing ten or more articles per year

Table 1.4a.is similar to Table 1.3a, but with the bar raised to ten articles per year (with no special provision for 2016). The total line is omitted from this and the next three tables since it's inherently identical to the total lines in the 1.3 tables.

Years > 9	Jan-Jn 2016	2015	Cum%	2014	2013	2012
None	4,041	5,723	2.2%	4,469	3,341	1,666
One	8,310	9,814	5.9%	5,394	4,402	2,964
Two	14,779	20,696	13.8%	8,867	6,081	3,638
Three	22,991	43,099	30.2%	28,468	8,932	4,022
Four	24,544	49,426	49.1%	44,543	26,006	11,030
Five	64,662	133,640		132,222	114,170	83,892

Table 1.4b. Articles in gray OA journals publishing more than nine each year

Years > 9	2015	Cum%	2014	2013	2012	2011
None	485	5.1%	556	528	480	424
One	632	11.7%	624	521	419	360
Two	971	21.9%	1,062	787	578	459
Three	1,018	32.6%	1,101	1,147	753	555
Four	1,329	46.5%	1,485	1,489	1,487	986
Five	5,096		5,096	5,096	5,096	5,096

Table 1.4c. *GOAJ* journals with more than nine articles per year

Years > 9	2015	Cum%	2014	2013	2012	2011
None	2,402	0.42%	2,750	2,618	2,597	2,200
One	8,866	1.99%	5,128	4,323	3,334	3,710
Two	30,263	7.33%	24,045	8,464	7,181	5,602
Three	40,084	14.40%	37,065	28,494	11,799	8,739
Four	57,394	24.52%	63,305	59,256	45,632	20,523
Five	427,913		427,743	390,320	368,101	319,575

Table 1.4d. Articles in GOAJ journals publishing at least ten articles per year

Revenues and Costs

While a much higher percentage of gray OA journals charge APCs, those fees are generally fairly low, as discussed in more detail in Chapter 6. As a result, although there were four-fifths as many 2015 articles in fee-charging gray journals as in fee-charging GOAJ journals, maximum total revenue was barely one fifth as much. Table 1.5 shows the details and can be compared to Table 1.3 in *GOAJ*.

	Jan-Jn 2016	**2015**	**2014**	**2013**	**2012**
Rev.	$46.419M	$81.130M	$64.551M	$51.087M	$37.713M
Pay art.	135,193	253,686	216,030	156,342	102,834
$/art	$343	$320	$299	$327	$367
Tot.art.	139,327	262,398	223,963	162,932	107,212
$/art	$333	$309	$288	$314	$352

Table 1.5. Revenue and cost per article by year

As in any revenue-related discussion, this table omits UA journals, since the APCs aren't known.

Starting Dates

Very few gray journals date back to the 20th century—not surprisingly since older journals should have either qualified for *DOAJ* or disappeared. But there's something considerably more interesting here, as shown in Figure 1.1, which *does* include UA journals.

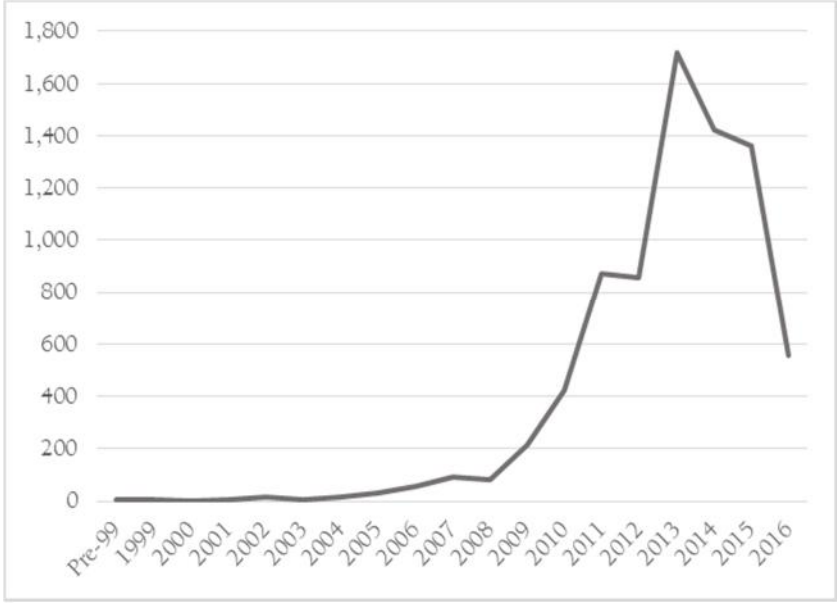

Figure 1.1. Gray OA journals by starting year

If you compare Figure 1.1 to Figure 1.1 in *GOAJ*, you'll see a much more rapid growth in recent years. The peak year for gray OA is 2013 with 1,717 journals, compared to the 953 new journals in *GOAJ* for 2011. These numbers include only journals that actually published articles: my sense is that template publishers "started" literally thousands of "journals" in 2013 and 2014, but that's another story.

Article Volume per Year, Gray and *GOAJ*

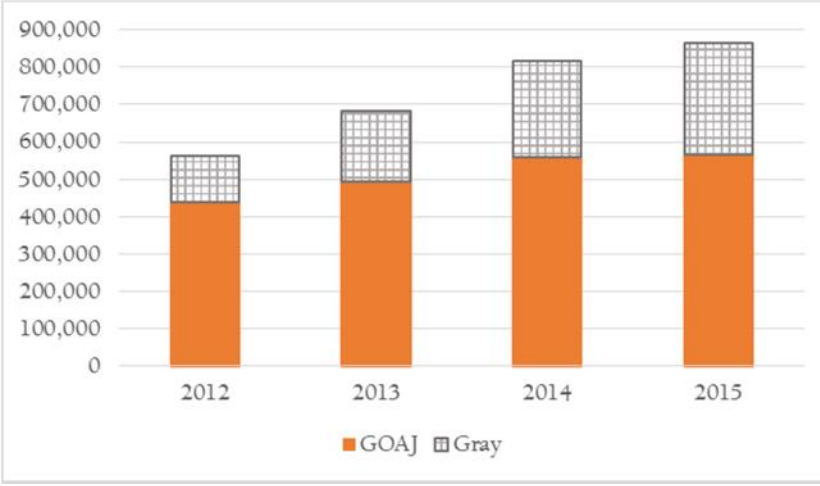

Table 1.2. *GOAj* and gray articles by year

Table 1.2 is *not* directly comparable to any table in *GOAJ*; instead, it compares overall article totals for *GOAJ* and gray OA on a year-by-year basis. The numbers appear in the total rows of Tables 1.3b and 1.3d.

Journal Growth and Shrinkage

Table 1.6 shows growth or shrinkage for gray journals (excluding UA) that published articles in 2014, 2015 or both. More journals shrank than grew. This table does *not* include journals with no articles in either year, although those could be considered "even" with no change.

Change 2014-2015	Count	Percent	Cum%
Grew 50%+	1,052	17.9%	
Grew 25-49.9%	294	5.0%	22.9%
Grew 10-24.8%	261	4.5%	27.4%
Even, ±9.99%	610	10.4%	37.8%
Shrank 10-24.9%	389	6.6%	44.4%
Shrank 25-49.9%	643	11.0%	55.4%
Shrank 50%+	1,311	22.4%	77.7%
No 2014 count	1,306	22.3%	
Total	5,865		

Table 1.6. Growth and shrinkage in gray OA journals

The Rest of This Report

The rest of this report goes into more detail about the journals and publishers of gray OA, although nowhere near as much detail as in *GOAJ*.

Chapter 2 discusses the very large number of "journals" that aren't counted, and includes some comparisons to *GOAJ*.

Chapter 3 peels the layers of the two source lists, specifically considering publishers and journals that aren't questionable OA at all. Chapter 3 also provides some comments on and measures of legitimately questionable journals.

Chapter 4 considers the Shen/Björk paper "'Predatory' open access: a longitudinal study of article volumes and market characteristics," specifically its estimate of "predatory" article counts, since those estimates have been used repeatedly as weapons against accepting OA.

Chapter 5 considers journals by article volume.

Chapter 6 considers fees and maximum revenues.

Chapter 7 discusses asserted country of publication.

Chapters 8-11 consider subjects and subject segments.

Chapter 12 is a thought experiment on gray OA without India.

The last chapter offers brief comments and conclusions for what is likely to be a one-shot study: as Shen/Björk say, "It would have taken a lot of effort to manually collect publication volumes and other data for all 11,873 journals"—and that was using the *much* smaller lists of 2014. I looked at some

20,000 websites, and it *was* a lot of effort; it certainly won't happen again without substantial sponsorship, and I'm not sure it's worth the effort. Meanwhile, that effort *has* happened, and these are the results.

2. Exclusions

Consider the journals with codes other than A or B—the ones excluded from most analysis (although one group, UA, is included in some discussions). Except for UA, articles in these journals weren't counted, in most cases because there was nothing to count.

While I'll discuss publishers in this chapter, there are publishers who show up in Chapter 3 and not in this chapter because they were excluded from consideration *as publishers*—e.g., because the publisher explicitly says it's a subscription publisher or because there's malware at the publisher level.

More Notes on Data Gathering

All visits were done in Chrome using either the Excel link-to-browser functionality or, for some publishers cleverly designing software such that it was impossible to use HTML copying and tricks to prepare a list of journals, directly within Chrome itself. I used Chrome because of the built-in language translation capabilities, although those rarely came into play: nearly all gray OA journals are in English.

Additional notes on the analysis, which began in July 2016 and ended in October 2016:

- If the URL didn't work, I stopped. Journals and publishers didn't get a second chance.
- For journals without clearly stated APCs or clear statements that there was no such fee, I assumed a hidden or missing APC (and assigned code UA) *unless* the journal was affiliated with a college, university, association or government or unless there was a clear statement of sponsorship.

- At all times, I used Malwarebytes Pro, Windows Defender and McAfee SiteAdvisor. If those tools (or, for that matter, Office itself) flagged the site as a security risk—either on its own or through outward links—I coded it as "XM" and stopped. I've been infected with some nasty malware twice in early studies of "journals" and wasn't taking any chances.
- I've come up with more ways to count articles in journals where the archives don't make it easy, but I gave up on 27 cases discussed below ("XO") and used approximations in some 90 cases—typically close approximations, e.g. determining the average number of articles per screen and counting screens or determining the average article length from one or two issues of a voluminous continuously-paginated journal and extrapolating the total count. In most approximation cases, 2016 and 2015 counts are *not* approximated.

Given exclusions and omissions, I believe the total count could be 10% or 20% higher than figures used here—or possibly up to 10% lower, if most of my approximations erred on the high side. I'd guess the deviation is much smaller, on the order of 2%-5%, but can't prove that.

The Codes—and a *GOAJ* Comparison

Code	Journals	% of Norm
UA: Unknown or hidden APC	902	22.4%
XE: Empty from 2012 through 2016	10,019	249.0%
XH: Hybrid	113	2.8%
XM: Malware	60	1.8%
XN: Not open access	135	3.4%
XO: Opaque, too difficult to count	72	1.8%
XU: Unworkable site	23	0.6%
XX: Unreachable or parking/ad page	746	18.5%
Total excluded	12,071	300.0%

Table 2.1. Journal exclusions for gray OA

Table 2.1 shows the fairly startling overall picture, discussed in more detail in the rest of this chapter. "% of Norm" is the number of journals as a percentage of what might be considered "normal" gray journals—namely, the 4,023 that have published five or more articles in at least two years and published at least one article in 2015.

Even without the huge number of empty "journals," most of which never had articles, ISSNs, editors or editorial boards or even brief descriptions, the excludable figures for gray OA are much higher than for *GOAJ*: nearly four times as many journals and roughly eight times the percentage of normal journals, 51.0% compared to 6.3%.

As an indication of just *how* startling the percentages are, Table 2.2 replicates Table 3.1 from *GOAJ* and adds a % of Norm column based on the norm for *GOAJ*: 9,250 journals. To make the tables fully comparable, add XP in Table 2.2 to XX, bringing that up to 240 or 2.6%, and XI to XO, bringing that up to 21 or 0.2%.

Code	Journals	% of Norm
CA: Missing or hidden APC	112	1.2%
XE: Empty 2011 through 2015	40	0.4%
XI: Impossible to count articles	15	0.2%
XM: Malware	103	1.1%
XN: Not open access	55	0.6%
XO: Opaque, too difficult to count	6	0.1%
XP: Parking or ad page	44	0.5%
XT: Translation inadequate	1	0.0%
XU: Unworkable site	37	0.4%
XV: Merged with no way to count	11	0.1%
XX: Unreachable	196	2.1%
Total excluded	620	6.7%

Table 2.2. Journal exclusions for *GOAJ*

UA: Unknown or Hidden APCs

I believe it *is* fair to describe these as predatory journals: the publisher asks the author to trust them that a "nominal" fee will indeed be reasonable. Even if a journal charges a range of APCs based on legitimate

variables, there's no excuse for failing to state the top of that range or the range itself.

Sixty-five publishers had UA journals and *no* AB journals, with a total of 663 journals among them, but two thirds of these had no more than five UA journals with actual articles. At the other end are a baker's dozen of publishers with 20 or more journals, all either UA, empty or excludable: Austin Publishing Group, International Digital Organization for Scientific Information (IDOSI), Lawarence Press, SM Group Open Access Journals, Universal Research Publications, ClinMed International Library, Priyanka Research Journal Publication, SciDoc Publishers, Insight Knowledge, Medwell Journals, SciRes Literature, Science Alert and Modern Scientific Press.

Among them, those 13 publishers account for 464 UA journals (in addition to 273 empty journals and 15 others); the other 52 all-UA publishers total 199 more. Others are in smaller or mixed publishers, including 102 "singletons."

XE: Empty from 2012 through 2016

Most of this enormous group is "journals," although there are some that faded away before 2012 and a few that may start publishing in the second half of 2016.

More than two-thirds of these are from another baker's dozen, this time a dozen template publishers and one publisher that lavished more care on its "journal" site than most template publishers. Here's the list—from most empty "journals" to fewest, although fewest in this case is still 345: Adyan Academic Press, British Open Research Publications, European Union Research Publishing, Eurasian Research Publishing, North American Research Publishing, Academic Knowledge and Research Publishing, Asian and American Research Publishing Group, American Research Publications, Canadian Research Publication, Academic and Scientific Publishing, International Organization of Scientific Research and Development (IOSRDD), Science and Technology Publishing and Research and Knowledge Publication. (One other template publisher, Sciences & Engineering Research Publication, had a mere 129 journals.) Those "publishers" account for 6,847 "journals."

With the exception of IOSRDD, these typically use one of three templates that appear to be identical other than journal names used—the same layout for journal lists, the same mostly-empty "journal" pages. If you look at the "publisher" names, you can see other similarities. The

sites typically list the same APC for all journals with different levels for different nations ($300 for wealthy nations seems most common). A journal's page may (or may not) have a templated "about" paragraph added if an article is ever submitted, possibly—rarely—also an ISSN or an editor/editorial board.

Some of these do land an occasional paper, but not many. I visited all the sites for most of these publishers. Although the 13 publishers listed above had a total of 272 journals with at least one article between 2012 and June 2016, "at least" was frequently also *at most:* I count a total of 524 articles over 4.5 years for the 272 journals, and no more than 250 in any year (2013, the peak year for these publishers: the total was 77 articles in 2015 and 25 in the first half of 2016).

There are quite a few other entirely-empty or mostly-empty publishers (I count 59 other publishers with at least 75% empty journals including 40 with at least 90% empty journals). With all of these mostly-empty publishers removed, we're down to 1,396 empty journals—still a lot, but only 29.8% of the gray OA norm.

It's tempting to include an essay on how to create an OA "publisher," since it's a process that would cost almost nothing and probably take less than a day's effort:

- Come up with an appropriate name—e.g. Beall Open Research Publications or Berkeley Research Publishing. Register the relevant domain.
- "Borrow" the template from one of the current template publishers—and for that matter you can probably "borrow" the journal list as well, since it's likely to consist of a standard prefix followed by a subject name and "Journal" or, in some cases, "Journal of" followed by a subject name. (Surprisingly, template publishers seem not to use "International Journal" all that often, although there are more than 1,100 empty journals beginning "International Journal of." Some template publishers vary the pattern.) If you're *really* ambitious, go for Adyan Academic Press—by far the largest set of "journals"—and change "Universal Open" to "Beall Open" or "Berkeley," but you're probably better off with something smaller like British Open (preface "British Open Journal of") or North American Open (preface "North American Open" with "Journal" at the end).

- Shazam: populate the pages and you're in business. A spam email to any of dozens of researchers should get you added to Beall's list, proving free publicity for your "journals."

How much redundancy is there in these templated journal titles? If you take all the empty journal titles for these 13 publishers and delete the common words (Journal, Research, Applied, Advances, Eurasian, American, British, Global, Universal, Open. North, Canadian and a couple of others) you have a list with—for example, and *not* including variations—11 Accounting, 12 Aerospace Engineering, 16 Agricultural (from 13 publishers), a total of 252 Agricultural or Agriculture including subtopics, 12 Analytical Chemistry, 19 Anthropology, 12 Archaeology, 13 Astronomy. That's just in the A's.

In all, there are 2,134 "unique" core titles and 4,713 duplicates—but "unique" includes, for example, counting "Addiction" and "Addiction and Therapy" as distinct titles.

You could spend a little money and not much time creating your own huge OA publisher—but it's a silly idea even if you lack ethics. Even at the full $300, none of these "publishers" could have taken in more than $9,300 in 2015 (more likely about one-third of that with most papers coming from low-income nations), and only four could have earned even $1,500. You're probably better off posting funny cat videos.

XH: Hybrid

Journals were flagged as hybrid either because the website explicitly called the journal hybrid or because current issues showed a mix of OA and subscription-only access. The 113 journals do *not* include journals from ten publishers self-identified as hybrid on the publisher's site.

While I didn't go searching for clues as to a publisher's country, I did note those that were fairly clear—and there's no getting around it: almost all of the XH journals are from India, at least 102 of the 113. All 102 are from Brainy Buzz, Literati Scientific and Publishers (Literati Publishers) and OMICS International.

XN: Not Open Access

Journals were flagged as not OA either because they label themselves as subscription, have embargos or require registration—or because at-

tempts to open articles were met with refusals of some sort or an inability to get from abstracts to full text. One single-journal publisher (with "Silicon Valley" in the name but openly based in India) requires registration even to see tables of contents!

The count here does not include 44 publishers clearly self-identified as subscription or not OA at the publisher level: those represent padding in the publisher list or part of an expansion from OA to whatever Jeffrey Beall doesn't like.

XO: Opaque, too difficult to count

I was pleased to reduce the *GOAJ* XO count to a mere 6 journals. I couldn't do as well here, not without spending (literally) hours on each journal. Combinations of "clever" programming and other problems were especially evident with Convergence Information Society, Council for Innovative Research and Institute of Research Engineers and Doctors (IRED); those three accounted for 49 of the 72 problematic journals.

XU and XX: Unworkable or Unreachable

I now believe the distinction between these two is arbitrary; think of them as totaling 769 journals that couldn't be reached or just didn't work—as compared to 277 for XP, XU and XX combined in *GOAJ*.

Publishers with lots of XU/XX journals include Academic World Education & Research Center, Access International Journals, APST Publication, Basic Research Journals, German Science and Technology Press, Horizon Journals, International Association for Engineering and Management Education (IAEME), Jacobs Publishers, Journal of The International Association of Advanced Technology and Science (JIAATS), Science Publishing Group and Signpost e Journals. Four of those have all XX journals—either 404 or parking pages or missing archives.

This doesn't include another 256 or so "publishers" that fail at the publisher level—most commonly yielding DNS or 404 errors or parking/ad pages.

Among journals, the most common problems include 404 errors (pages do not exist—265 of them), "journals" that are now ad or parking pages or suspended accounts (59), database errors (18, all from one publisher) and DNS lookup failures (254). There are also a range of other problems including unresponsive pages, lack of archives and failure to ever finish loading PDFs.

3. Breaking Down the Lists and Questionable Journals

Beall's publisher and journal lists have grown rapidly, and that growth has been widely publicized. Since those lists are the basis for this report, it makes sense to look at them a bit.

Publishers

Reason	Count
XX: Unreachable/unworkable	257
XN: Not an OA publisher	44
XM: Malware at publisher site	39
Duplicates another publisher	21
All journals in DOAJ	19
XH: Publisher-level hybrid statement	9
XO: Too difficult to unravel	5
FP: Entirely obvious plagiarism	1
Total	395

Table 3.1. Publishers not included in gray OA

Table 3.1 shows reasons why nearly 40% of the 1,025 "publishers" in the Beall list as of early July 2016 (after eliminating absolute duplicates) weren't evaluated further, arranged from most common to least common reasons. (Another 34 publishers had entirely empty journals,

and 12 others had entirely XU/XX journals, but those are included in other discussions.)

The XN cases are especially interesting as they represent publishers that clearly state that they don't publish OA journals (or, as with some XX cases, aren't journal publishers at all).

The single "FP" case didn't get an X code because it was a single case. It's a "publisher" where each "journal" had a single "issue" (in 2014, I think)—and each issue had a single article. When I did title searches for the article titles, they showed up in other journals. In other words, the "publisher" was a pathetic attempt to attract new authors by salting the journals with existing papers—an attempt that utterly failed, since there were no other papers at all.

It's hard to know what to say about the huge number of XX "publishers," most of which yielded DNS or 404 errors or ad/parking pages. To the extent that they ever had journals with articles, those articles may now be stranded—but they may also have been attempts that never actually yielded any published articles.

Evidence

Going through the entire *Scholarly Open Access* archive through June 2016, I found only 112 publishers where Beall had made even a moderate case. ("The publisher has a funny name" and "I think these subjects have enough journals" are *not* cases.) That's 112 out of 1,025.

Of the 112 plausibly questionable publishers, one was a duplicate, three had entirely empty journals, eight had malware, seven weren't OA publishers at all, one was obscure, and 29 couldn't be reached. That leaves 62 plausibly questionable publishers, accounting for some 3,600 journals (including some 1,900 A and B) and around 63,000 articles in 2015.

Questionable for Other Reasons

But there are other publishers (and journals) that are fairly clearly questionable, even without qualitative analysis:

- The already-noted "publisher" with entirely plagiarized articles.
- Sixty-one publishers where all journals either lacked APCs, were empty or had X codes. That is, none of these were "good" journals.

These and other journal-level cases show up in the questionable journal analysis, which follows.

"Singleton" Journals

Reason	Code
XX: Unreachable/unworkable	170
XN: Not OA	52
XM: Malware	34
XO: Obscure/uncountable	16
XE: Empty	13
XH: Hybrid	8
Total	293

Table 3.2. Singleton journals not fully analyzed

Table3.2 shows reasons that 293 of the 900 journals aren't fully analyzed. Aanother 102 have missing or hidden APCs, leaving 506 coded A or B. Of those 506, only 19 have plausible cases made against them in Beall's posts.

Questionable and Predatory: The Broadest View

It's clear that Jeffrey Beall expects people to just trust what he says in the vast majority of entries in his lists. It's also clear that the publisher list goes far beyond OA and, indeed, beyond journal publishers.

That said, Beall does make a case against some of the largest gray OA publishers—and there are fairly clear questionable cases beyond those where he's made a case.

In one sense, every journal in this study (except those founded in 2016) is somewhat questionable, the reasonable question being "Why isn't it in *DOAJ*?" But in doing the quantitative study here, I couldn't help but notice some qualitative issues along the way. I flagged some journals as being clearly questionable (albeit without a Beall case) for five reasons:

- **A: APC hidden or missing**. Already discussed, these "UA" journals are not just questionable, they're predatory.
- **B: Beall makes a case**.
- **C: Crackpottery**. A handful of journals, mostly with physics in the title, seem to feature papers that mathematically disprove Einstein's theories or otherwise seem on the fringe. (On the other hand, claims of arsenic-based life appeared in a highly-regarded non-OA journal, Beall was fond of trashing journals for papers linking glyphosate to cancer until the World Health Organization supported that claim, and articles suggesting tectonic plates were probably regarded as crackpottery in the early 20th century, so I wouldn't push this one too hard).
- **L: Loremipsum in page**. Journal sites that actually have paragraphs of loremipsum text or other nonsense text where vital information should be.
- **P**: **Papermill**. Journals that show clear evidence of publishing random articles with absurdly short review periods.
- **S: Single author**. This special category is the "Eluozo category"—S.N. Eluozo, a Nigerian scholar who published three articles in a single issue of each of 16 or more journals in 2013 or 2014, almost always the only articles ever published in those journals (from a template publisher). From what I can see, the articles are all legitimate science but very narrow—and most good journals do try to publish more than one author. Many don't allow multiple papers from an author within an issue.

Some journals belong in more than one category. Generally, B takes precedence, followed by A, followed by others—thus, a papermill with hidden APCs is coded A, not P.

An important caveat here: Good papers appear in questionable journals—and questionable publishers are as likely to have good journals as renowned publishers are to have fake journals or journals devoted to more-than-questionable science. If I had to guess, I'd guess that the bulk of articles in the tables that follow are legitimate scholarship and research, frequently in narrow fields. However, it's also fair to suggest that papermills, almost all of which hail from India, are really *certificate* mills: the authors need the certificates of publication to fulfill university requirements.

Category	Jan-Jn2016	2015	2014	2013	2012
B: Beall evidence	1,533	1,693	1,594	1,382	829
A: APC missing/hidden	424	529	439	408	295
P: Papermill	70	72	61	40	26
L: Loremipsum text	8	7	10	8	
C: Crackpottery	5	4	5	5	4
S: Single author	1		6	16	
Questionable sub	**2,041**	**2,305**	**2,115**	**1,859**	**1,154**
Others	3,172	3,683	3,072	2,324	1,400
Total	**5,213**	**5,988**	**5,187**	**4,183**	**2,554**
Questionable %	39.2%	38.5%	40.8%	44.4%	45.2%

Table 3.3. Questionable journals

Category	Jan-Jn2016	2015	2014	2013	2012
B	33,848	70,074	62,191	44,376	30,010
A	14,351	30,955	26,019	22,916	15,753
P	28,281	46,892	31,365	22,161	9,925
L	17	27	39	22	
C	202	451	499	438	276
S	1		18	53	
Quest.	**76,700**	**148,399**	**120,131**	**89,966**	**55,964**
Others	78,647	148,564	135,052	98,679	69,075
Total	**155,347**	**296,963**	**255,183**	**188,645**	**125,039**
Quest. %	49.4%	50.0%	47.1%	47.7%	44.8%

Table 3.4. Questionable articles

Tables 3.3 and 3.4 summarize the situation, and they're fairly revealing. (Table 3.3 is in descending order by number of journals publishing articles in 2015; Table 3.4 uses the same order for consistency.)

Among other things, it's worth noting that—while legitimately questionable journals publish roughly half of gray OA articles—cases where Beall made a legitimate case accounted for only half of questionable cases, less than one-quarter of all gray OA articles, and only one-eighth

as many articles as in *DOAJ* journals in 2015. Also noteworthy: there aren't a lot of papermill journals but they churn out a lot of articles, as you'd expect. Finally, the three smaller questionable categories are so small they might not be worth mentioning, never totaling even 600 articles in a year.

4. The Shen/Björk paper

Cenyu Shen and Bo-Christer Björk published "'Predatory' open access: a longitudinal study of article volumes and market characteristics" in *BMC Medicine* 13, October 2015. (I'm bemused at the idea that this is a medical paper, but that's a separate discussion.) I started questioning the paper's conclusions as soon as it appeared, and continued to do so in my blog and in *Cites & Insights*.

Quite apart from the apparent assumption that Beall's word is gospel when it comes to journals being "predatory"—an assumption I found, and find, appalling—I thought the numbers were implausible. The authors used a sample of 613 journals to assert that there were around 8,000 active "predatory" journals in 2014 and that those journals published around 420,000 articles in 2014 (up from around 310,000 in 2013 and 212,000 in 2012).

Being presented with a case for the implausibility of the numbers, the authors responded that the article was peer-reviewed and used proper statistical methods. As I was writing this, I took the time to read open reviewer comments on the article and the authors' responses. Notably, *all* of the reviewers said they weren't qualified to review the statistics—and there were certainly questions raised about the assumption that to be on Beall's list was to be predatory.

The authors are right about one thing: looking at all the journals is a ridiculously large task. But that task showed that gray journals are just as heterogeneous as I thought they were, making it easy for a 6% sample to be wildly off base.

The First Cut

Now that I've done the work, the first note could be that the article's 2014 figure has the first two digits reversed: it's closer to 240,000 than to 420,000. Of course, the authors did not accidentally transpose digits; they came up with too-large results. Instead of 420,000 for 2014, 310,000 for 2013 and 212,000 for 2012, the figures should be 255,000 for 2014, 189,000 for 2013 and 125,000 for 2012 (rounding to the nearest thousand)—consistently between 59% and 61% of the article's figures.

"255,000 questionable as compared to 560,000 *DOAJ*" isn't as astonishing as "nearly as many predatory as not." That 420,000 figure has been cited a *lot*, mostly by critics of open access in general.

But there's more to say…

The Second Cut

The authors were working from an earlier and much smaller pair of Beall lists than those that I worked from. I used the Wayback Machine to download versions of the list as close as possible to the versions they used (in both cases, later and presumably a little larger). Flagging publisher and journal listings from those earlier versions yield the figures in Table 4.1, including "UA" journals but excluding X-coded ones.

	2014	2013	2012
Journals	2,692	2,222	1,370
Articles	113,996	87,325	55,303

Table 4.1. Journals and articles based on Beall lists at time of Shen/ Björk article

Now we're down from 8,000 active journals to 2,692—and from 420,000 articles to just under 114,000. The percentages are still clustered: now the real numbers are 26% to 28% of those reported in the article. Even if you added 50% to my figures to account for a few dozen not-fully-counted journals (rather than the 5% to 10% I consider plausible), you'd be nowhere near 200,000, let alone 420,000. And, of course, 114,000 is a pretty small fraction of 560,000—just over one-fifth.

Even those numbers involve the odd assumption that Beall's tagging is definitive. What happens if we reduce the universe to those articles and publishers where Beall's actually made a case?

The Final Cut

	2014	2013	2012
Journals	936	781	488
Articles	29,947	21,500	13,198

Table 4.2. Journals and articles where Beall made a case

Table 4.2 shows the results: fewer than 30,000 articles in 2014—about 7% of the article's estimate. (The 2012 and 2013 figures are 6% to 7% of the article's estimates.) These are cases where Beall not only listed a publisher or journal at the time the authors downloaded the lists, but actually made a case for the journals or publishers being questionable or "predatory."

Those numbers are too low—but they're arguably what should have emerged from the study. As noted in Chapter 3, I believe realistic numbers are on the order of 120,000 for 2014; 90,000 for 2013; and 56,000 for 2012—still a *lot* of articles appearing in questionable journals, but not quite so alarmingly high.

What Went Wrong?

How could these two scholars be so far off? First there's the assertion that all journals on Beall's lists are actually predatory. Second, the "stratified" random sampling method involves some tricky assumptions, based on a "suspicion" that was "verified" by sampling all of ten journals—the suspicion "that journals from small publishers often publish a much higher number of articles than those of large publishers."

The sampling used in this study yielded a much lower percentage of empty journals than my 100% survey. The article estimates that 67% of listings represent active journals; my 100% survey (admittedly of a larger list) shows 40% active journals. That's an *enormous* difference: instead of 8,000 active journals from the smaller list, you wind up with around 4,800. That's probably about right (I show 5,988—but that's from a much larger list).

Beyond that, it appears that the sheer heterogeneity of journals makes projection from a small sample so dicey as to be useless. Unfortunately, I believe that to be the case.

5. Article Volume

This is the last chapter to include journals with missing or hidden APCs; the rest of the report, including part of this chapter, includes only journals coded A or B.

Most gray journals don't publish very many articles, although there are exceptions. Two journals published more than 4,000 articles in their peak year (2012-2016), one more published more than 3,000 and 11 more broke the 2,000-article mark—but only two journals published 2,000 articles in each of three years, and *none* managed that level in each of four years.

Consider three ways of breaking down article volume: ten groups based on roughly equal numbers of journals, ten groups based on roughly geometric doubling, and the quintiles used in *GOAJ*.

Roughly Equal Journal Numbers

Table 5.1. breaks down gray journals (excluding X codes) into ten roughly equal parts—"roughly" because 774-journal boundaries almost always occur within a run of journals with the same peak number of articles.

Note "peak number" here and throughout this chapter: the highest number of articles during the 4.5 years. As the table makes clear, it is *never* the case that all journals within a size range published articles in any given year. Closest are the largest journals (97 or more articles), and even there at least 14 of 766 journals were wholly absent in any given year. The worst case is the lowest *and largest* group, 991 journals that never published more than two articles per year: no more than 364 of these, 37%, published in any given year.

Articles	Journals	Jan-Jn2016	2015	2014	2013	2012	Cum%
97+	766	737	752	720	620	481	10%
49 to 96	773	717	745	696	591	450	20%
31 to 49	744	648	699	657	552	380	29%
20 to 30	850	684	770	698	572	348	40%
14 to 19	755	543	657	544	441	226	50%
10 to 13	750	504	623	492	347	165	60%
7 to 9	727	428	560	459	325	163	69%
5 or 6	731	414	456	317	217	108	79%
3 or 4	656	272	362	287	226	124	87%
1 or 2	991	266	364	317	292	109	

Table 5.1. Journals grouped by size that published articles each year.

The **Cum%** column in Table 5.1 shows how close I could come to 10% groupings while respecting whole-number boundaries: ideally, every percentage would end in zero. Contrast that to Table 5.2, which shows article totals year by year and cumulative percentages for 2015, the most recent full year: more than two-thirds of all 2015 articles are in the 10% most prolific journals, and the bottom 40% of journals account for only 1.2% of articles

Articles	Jan-Jn2016	2015	2014	2013	2012	Cum15%
97+	106,446	206,959	179,627	129,300	84,372	69.7%
49 to 96	20,105	38,291	32,016	25,189	19,404	82.6%
31 to 49	10,265	19,503	17,074	13,495	9,533	89.2%
20 to 30	6,764	13,136	11,779	9,077	5,867	93.6%
14 to 19	4,166	7,306	5,814	5,037	2,578	96.0%
10 to 13	2,987	5,148	3,870	2,799	1,357	97.8%
7 to 9	1,952	3,165	2,552	1,854	962	98.8%
5 or 6	1,617	1,955	1,274	881	456	99.5%
3 or 4	701	1,003	760	628	352	99.8%
1 or 2	344	497	417	385	158	

Table 5.2. Articles in journals grouped in ten roughly equal parts

Halves: Journals Grouped in Logical Groupings

Articles	Journals	Jan-Jn2016	2015	2014	2013	2012	Cum%
1,000+	54	53	54	53	48	41	1%
500-999	73	71	73	72	58	44	2%
250-499	150	148	149	141	121	99	4%
125-249	344	327	338	324	280	210	8%
63-124	598	563	577	545	465	355	16%
32-62	1,007	894	954	889	753	537	29%
16-31	1,357	1,052	1,219	1,083	882	521	46%
8-15	1,520	1,003	1,239	1,002	741	357	66%
4-7	1,295	706	834	616	421	216	83%
1-3	1,345	396	551	462	414	174	

Table 5.3. Journals by peak volume, logical groupings

Table 5.3 groups journals by logical groupings—starting with 1,000+ and going to roughly half the number for each lower group. I find it interesting that the cumulative percentages are also roughly inverted for the first five rows, with cumulative percentage doubling in each row.

Articles	Jan-Jn2016	2015	2014	2013	2012	Cum15%
1,000+	39,361	77,494	60,583	45,693	41	26.1%
500-999	17,657	35,851	34,374	22,150	44	38.2%
250-499	20,744	38,446	35,012	23,641	99	51.1%
125-249	22,650	44,229	39,705	30,121	210	66.0%
63-124	19,148	36,751	31,590	24,683	355	78.4%
32-62	16,599	30,994	26,407	20,848	537	88.8%
16-31	10,004	18,846	16,656	13,180	521	95.2%
8-15	5,899	9,825	7,670	5,899	357	98.5%
4-7	2,643	3,565	2,432	1,749	216	99.7%
1-3	642	962	754	681	174	

Table 5.4. Articles for journals grouped by logical groupings

The 54 largest journals account for 26% of 2015 articles; the largest 4% for more than half. That's really no surprise.

Journals and Articles by Segment

The remaining tables use the same size groupings as in *GOAJ* and do *not* include UA (hidden or no APC): I didn't assign subjects or segments to those journals. Tables 5.5 and 5.6 are directly comparable to Tables 4.4 and 4.5 in *GOAJ*, except that Tables 5.5 and 5.6 lack free %. *GOAJ* labels the rows Largest, Large, Medium, Small and Smallest; I've omitted those labels and added percentage rows and columns.

	HSS	Biomed	STEM	Total	%
600+	35	16	33	84	2%
150-599	128	69	172	369	7%
60-149	192	179	279	650	12%
20-59	367	506	646	1,519	29%
1-19	547	1,046	1,037	2,630	50%
Total	1,269	1,816	2,167	5,252	
%	24%	35%	41%		

Table 5.5. Journals by segment, 2015

Table 5.5 includes only A and B-coded journals that published articles in 2015.

	HSS	Biomed	STEM	Total	%
600+	38,335	16,334	35,503	90,172	34%
150-599	27,054	13,424	34,039	74,517	28%
60-149	12,934	12,673	17,596	43,203	16%
20-59	9,588	13,919	14,380	37,887	14%
1-19	3,839	6,468	6,312	16,619	6%
Total	91,750	62,818	107,830	262,398	
%	35%	24%	41%		
Art/jrnl	72	35	50	50	

Table 5.6. Articles by segment, 2015

Table 5.6 shows 2015 article counts by size of journal and segment—but also the average articles per journal. Notably, humanities and social sciences (HSS) represent a much lower percentage of gray OA than of *GOAJ*—and, paradoxically, HSS journals have the *most* articles per journal and Biomed the fewest, with Science, Technology, Engineering and Math (STEM) in the middle. (In *GOAJ*, the 2015 averages are 30 for HSS, 77 for biomed and 75 for STEM.)

There's a simple explanation for the high average articles for HSS: 51,219 of the 91,750 articles are in 220 journals I tagged as "Miscellany" because they covered so many subjects, including some within HSS. That's 75 more journals and roughly 40,000 more articles than in *GOAJ*'s Miscellany group.

6. Fees and Maximum Revenue

More than 90% of gray OA journals charge fees—but most of them don't rake in big bucks. That's considerably truer for gray OA journals than for the minority of *GOAJ* journals that charge fees. This chapter uses essentially the same table formats and limits as Chapter 5 of *GOAJ* to provide some comparability—and like Chapter 5 of *GOAJ*, uses 2015 figures.

Revenue Ranges

The basis for calculating APCs and maximum potential 2015 revenue is the same as for *GOAJ*. The APC is the fee stated for an American (or "wealthy nation") author of a 10-page full research paper; if society members are offered discounts (and membership is not required), I use the non-member price; if students and other special categories have discounts, I use the most expensive price category.

The revenue figures are almost certainly much too high for most journals, since they don't include waivers or lower prices for less wealthy nations. Based on casual observations, I'd guess that *most* articles are published at lower-income-nation prices.

If you compare Table 6.1 with Table 5.1 in *GOAJ* you'll see startling differences. Although there are more than twice as many gray OA journals with known fees as there are *GOAJ* journals with known fees, only 69 gray journals could have taken in $150,000 or more in 2015, compared to 397 *GOAJ* journals—and to get past the 1,000-journal mark, which gets down to $30,000 for *GOAJ*, means going down to $10,000 for gray OA. At the very bottom, 428 *GOAJ* journals either took in less than $1,000 or didn't publish any 2015 articles—whereas the figure for gray OA is 2,611 journals.

Revenue	Journals	Cum J	Articles	Art/J
$9 to $10 million	1	1	3,356	3,356
$2 to $8.9 million	0	1	0	
$1 to $1.96 million	3	4	7,941	2,647
$750,000 to $999,999	1	5	2,250	2,250
$500,000 to $749,999	3	8	4,305	1,435
$400,000 to $499,999	4	12	1,360	340
$300,000 to $399,999	4	16	1,547	387
$250,000 to $299,999	10	26	3,043	304
$200,000 to $249,999	11	37	7,153	650
$150,000 to $199,999	32	69	22,712	641
$100,000 to $149,999	55	124	17,802	324
$75,000 to $99,999	83	207	18,289	220
$50,000 to $74,999	114	321	23,032	202
$40,000 to $49,999	72	393	8,839	123
$30,000 to $39,999	122	515	15,011	123
$25,000 to $29,999	85	600	11,288	133
$20,000 to $24,999	140	740	13,518	97
$15,000 to $19,999	201	941	16,345	81
$10,000 to $14,999	314	1,255	17,270	55
$7,500 to $9,999	282	1,537	12,250	43
$5,000 to $7,499	463	2,000	13,621	29
$2,500 to $4,999	821	2,821	17,100	21
$1,000 to $2,499	942	3,763	11,133	12
$1 to $999	1,105	4,868	4,521	4
$0 (no 2015 articles)	1,506	6,374	0	

Table 6.1. Revenue by journal, detailed breakdown

Detailed APC Breakdown

APC	Journals	Cum J	Articles	Art/J
$3,619	1	1	25	25
$2,500-$2,900	22	23	4,830	220
$2,250-$2,319	3	26	246	82
$2,000-$2,119	51	77	3,476	68
$1,800-$1,949	61	138	1,382	23
$1,500-$1,735	91	229	2,440	27
$1,250-$1,438	60	289	770	13
$1,000-$1,249	226	515	4,592	20
$750-$999	372	887	8,333	22
$600-$749	266	1,153	4,932	19
$400-$599	1,075	2,228	25,047	23
$300-$399	1,208	3,436	18,905	16
$200-$299	843	4,279	21,465	25
$100-$195	1,045	5,324	64,025	61
$10-$99	1,049	6,373	90,985	87

Table 6.2. APC levels, detailed breakdown

Table 6.2 uses the same breakdown as Table 5.2 in *GOAJ*—except that there are no gray OA journals with APCs greater than $3,619 and limits on other rows are adjusted for reality (e.g., there are no $2 APCs in gray OA). As with the rest of this chapter, one or two now-you-see-them, now-you-don't journals are missing from these figures.

Fees and Revenue by Segment

Table 6.3 uses the same categories and layout as Table 5.4 in *GOAJ* to show 2015 values for active journals, APC level, article counts and maximum potential revenue by broad subject segment. The most striking group is probably the top one, where potential revenues for biomed articles in the most expensive journals are more than times the combined total for HSS and STEM.

The much higher revenue figures for HSS than for HSS in *GOAJ* are largely due to miscellaneous journals including papermills, and it's probably worth repeating that *actual* revenues are probably much lower than maximum revenues, quite possibly less than half as much.

	HSS	Biomed	STEM
$1,400+	3	166	34
Articles	132	11,173	1,254
Revenue	$218,457	$25,263,734	$2,247,256
$600-$1,399	48	464	184
Articles	2,291	9,645	6,531
Revenue	$1,990,572	$9,170,083	$5,752,026
$200-$599	519	678	1,003
Articles	28,769	10,512	26,136
Revenue	$9,400,038	$4,240,630	$9,042,431
$10-$199	615	365	789
Articles	57,440	29,054	70,749
Revenue	$5,184,227	$2,428,885	$6,192,007
Free	84	143	157
Articles	3,118	2,434	3,160

Table 6.3. Articles and revenues by segment

There may not be a lot more to say here. Most gray OA journals, no matter what the subject, charge—but generally charge moderate fees.

7. Country of Publication

This chapter covers only a subset of A- and B-coded gray OA, and it's a defective subset at that. To wit:

- This study is primarily descriptive, not investigative: I did not go beyond the websites themselves looking for country of publication. For 28% of the journals active in 2015 (25% of the 2015 articles), I did not record a country of publication at all.
- I accepted what was stated at face value—with one key exception: if two contact points or offices in two different countries were provided, and if the first was in the United States, United Kingdom or Canada and the second was not, I looked at the language on the website. If it was clearly not typical of native English syntax, I recorded the other country as the country of publication. (A helpful hint: "Copyright" is a single word in the US, UK and Canada. There are other dead giveaways, but that one is readily avoidable.)
- But consider the first eight words of the bullet above. I would guess that 90% or more of the journals listed as being published in the United States, United Kingdom or Canada are actually published elsewhere, based on the peculiar syntax of the webpages.

What may be most interesting about the tables that follow is what's *not* there: namely, gray OA isn't a major phenomenon in all of the global South. To wit, there are *no* gray OA journals coded A or B in South America and only five in all of Latin America (four in Mexico and one in the British Virgin Islands). For that matter, although overall this subset has 45% as many 2015-active journals as *GOAJ*, with 39% as many articles, those percentages are both below 10% in Eastern Europe and 20% in the Middle East and Western Europe.

Country	Journals	%Free	Articles	%Free
Algeria	1	100%	18	100%
Australia	34	12%	1,255	31%
Austria	10	0%	1,152	0%
Bangladesh	22	9%	2,613	0%
Bosnia and Herzegovina	4	50%	254	64%
British Virgin Islands	1	0%	521	0%
Bulgaria	28	14%	1,824	2%
Canada	153	7%	7,892	2%
China	8	63%	266	73%
Croatia	5	0%	374	0%
Cyprus	1	100%	3	100%
Czech Republic	2	50%	88	86%
Denmark	1	0%	127	0%
Egypt	11	9%	176	1%
France	1	0%	58	0%
Georgia	1	0%	13	0%
Germany	12	92%	216	72%
Ghana	3	0%	34	0%
Hong Kong	58	48%	614	31%
Hungary	1	0%	93	0%
India	2,033	5%	161,910	1%
Indonesia	4	0%	78	0%
Iran	15	40%	727	18%
Iraq	1	100%	49	100%
Italy	2	0%	207	0%
Japan	5	0%	1,131	0%
Kazakhstan	1	0%	12	0%
Kenya	15	0%	254	0%

Table 7.1a. Countries with gray OA journals, part 1

Country	Journals	%Free	Articles	%Free
Libya	1	0%	51	0%
Lithuania	7	0%	157	0%
Macedonia	7	43%	96	7%
Malaysia	44	20%	746	27%
Mexico	4	0%	123	0%
Mongolia	8	0%	58	0%
Morocco	3	33%	1,191	35%
Nepal	2	50%	123	5%
Netherlands	1	100%	3	100%
Nigeria	403	0%	2,779	0%
Pakistan	150	6%	3,925	3%
Philippines	3	100%	134	100%
Poland	1	0%	220	0%
Romania	29	3%	1,729	1%
Russia	2	0%	476	0%
Saudi Arabia	1	0%	72	0%
Singapore	20	0%	1,389	0%
Slovakia	4	0%	82	0%
South Korea	14	43%	1,014	4%
Sweden	1	0%	28	0%
Switzerland	14	43%	963	7%
Tunisia	10	10%	217	46%
Turkey	47	49%	1,307	55%
Ukraine	1	0%	10	0%
United Arab Emirates	36	8%	711	11%
United Kingdom	201	17%	5,800	11%
United States	853	4%	17,442	4%

Table 7.1b. Countries with gray OA journals, part 2

Country	Journals	%Free	GOAJ	Gray%
India	2,033	5%	461	441%
United States	853	4%	952	90%
Nigeria	403	0%	28	1439%
United Kingdom	201	17%	300	67%
Canada	153	7%	199	77%
Pakistan	150	6%	70	214%
Hong Kong	58	48%	39	149%
Turkey	47	49%	295	16%
Malaysia	44	20%	63	70%
United Arab Emirates	35	6%	14	250%
Australia	34	12%	114	30%
Romania	29	3%	322	9%
Bulgaria	28	14%	34	82%
Bangladesh	22	9%	31	71%
Singapore	20	0%	28	71%
Iran	15	40%	297	5%
Kenya	15	0%	7	214%
South Korea	14	43%	40	35%
Switzerland	14	43%	43	33%
Germany	12	92%	246	5%
Egypt	11	9%	16	69%
Austria	10	0%	50	20%
Tunisia	10	10%	1	1000%
China	8	63%	47	17%
Mongolia	8	0%	0	
Lithuania	7	0%	35	20%
Macedonia	7	43%	19	37%
Croatia	5	0%	103	5%
Japan	5	0%	94	5%

Table 7.2. Countries with five or more gray journals, in journal order

Country	Articles	%Free	GOAJ	Gray%
India	161,920	1%	54,650	296%
United States	17,442	4%	44,881	39%
Canada	7,892	2%	6,175	128%
United Kingdom	5,800	11%	23,098	25%
Pakistan	3,925	3%	5,833	67%
Nigeria	2,779	0%	1,965	141%
Bangladesh	2,613	0%	1,278	204%
Bulgaria	1,824	2%	1,479	123%
Romania	1,729	1%	12,734	14%
Singapore	1,389	0%	2,248	62%
Turkey	1,307	55%	13,838	9%
Australia	1,255	31%	3,190	39%
Morocco	1,191	35%	971	123%
Austria	1,152	0%	1,297	89%
Japan	1,131	0%	6,907	16%
South Korea	787	6%	5,106	15%
Switzerland	963	7%	2,282	42%
Malaysia	746	27%	3,419	22%
Iran	727	18%	13,621	5%
United Arab Emirates	711	11%	823	86%
Hong Kong	614	31%	3,390	18%
British Virgin Islands	521	0%	6	8683%
Russia	476	0%	10,625	4%
Croatia	374	0%	3,022	12%
China	266	73%	9,039	3%
Bosnia and Herzegovina	254	64%	290	88%
Kenya	254	0%	87	292%
Poland	220	0%	12,389	2%
Tunisia	217	46%	12	1808%
Germany	216	72%	12,218	2%
Italy	207	0%	10,855	2%

Table 7.3. Countries with 200 or more gray articles in 2015

Table 7.2 (limited to countries with at least five journals to fit the table on a single page) adds the number of *GOAJ* journals active in 2015 and gray as percentage of *GOAJ*, 45% overall but with very high percentages for Nigeria, India, UAE, Pakistan, Kenya and Hong Kong, all with more

gray journals than *GOAJ* journals. (Neither Tunisia nor Mongolia have enough *GOAJ* journals for a meaningful comparison.)

Number of journals, even limited to those actually publishing articles in 2015, can be a misleading figure, as may be evident when comparing Table 7.2 with Table 7.3. The percentages for India and Nigeria, while still having more gray articles than *GOAJ* articles, are much lower than for journals, for example.

Country	**Articles**	**%Free**
China	266	73%
Germany	216	72%
Bosnia and Herzegovina	254	64%
Turkey	1,307	55%
Tunisia	217	46%
Morocco	1,191	35%
Hong Kong	614	31%
Australia	1,255	31%
Malaysia	746	27%
Iran	727	18%
United Kingdom	5,800	11%
United Arab Emirates	711	11%
Switzerland	963	7%
South Korea	1,014	4%
United States	17,442	4%
Pakistan	3,925	3%
Canada	7,892	2%
Bulgaria	1,824	2%
India	161,910	1%
Romania	1,729	1%

Table 7.4. Countries with some non-APC journals and at least 200 2015 articles

Table 7.4 shows the relatively few countries with at least 200 gray OA articles in 2015 where *any* of the articles appeared in journals with no APCs, appearing in descending order be free (non-APC) percentage.

Regions

With India accounting for 93% of all Asian gray OA articles in 2015 and Nigeria accounting for 61% of all African articles (and fewer than 1,800 other articles), there's little point in detailed regional discussions.

Region	Journals	%Free	GOAJ	Gray%
Africa	436	1%	134	325%
Asia	2,371	7%	1,133	209%
Eastern Europe	94	13%	1,398	7%
Latin America	5	0%	1,783	0%
Middle East	111	31%	642	17%
Pacific/English	1,040	5%	1,204	86%
Western Europe	243	22%	1,935	13%

Table 7.5. Gray journals active in 2015 by region

Region	Articles	%Free	GOAJ	Gray%
Africa	4,544	12%	7,731	59%
Asia	174,001	2%	101,276	172%
Eastern Europe	5,431	6%	61,126	9%
Latin America	644	0%	65,298	1%
Middle East	3,042	32%	31,237	10%
Pacific/English	26,589	5%	55,372	48%
Western Europe	8,554	10%	81,267	11%

Table 7.6. Gray 2015 articles by region

Tables 7.5 and 7.6 show the regional summaries (noting that Pacific/English is probably several times too high and Western Europe is somewhat too high) compared with *GOAJ* 2015 numbers.

8. Subjects and Segments

When preparing *GOAJ*, I was able to assign subjects based primarily on the narrower subjects and keywords provided by publishers in *DOAJ*. The set of 28 subjects in three segments first appeared in *Open-Access Journals: Idealism and Opportunism* (ALA, 2015). As I said in *GOAJ:*

- Assignment of journals to one of 28 subjects is tricky and partly subjective.
- Assignment of subjects to segments may also be arguable, at least in the cases of anthropology and psychology, which some might argue belong in STEM and biomed respectively.

The first bullet is even truer this time around, since I based subject assignment on journal titles and article titles in recent issues, but primarily on journal titles.

Tables 8.1 through 8.3 show all 28 subjects; all subject and segment discussions cover A- and B-coded journals.

Table 8.1 shows all journals, the percentage of non-APC journals, journals active in 2015 and articles in 2015; it's in order by country.

Table 8.2 repeats the 2015 journal and article counts, this time in article count order, and adds an articles-per-journal figure.

Table 8.3 repeats the 2015 article counts, adds *GOAJ* 2015 article counts, and shows the gray count as a percentage of the *GOAJ* count: the overall percentage is 30.3%.

Note that "Other Sciences" includes (and is mostly) interdisciplinary/multidisciplinary journals that appear to be almost entirely in biomed and STEM, while "Miscellany" is mostly multidisciplinary journals that include a fair number of HSS articles. Most papermills ae in one of these two categories. Note also that "Economics" includes most business and management topics.

Subject	Journals	%Free	Active15	Articles
Agriculture	393	2.8%	305	8,756
Anthropology	63	7.9%	47	1,614
Arts & Architecture	40	7.5%	23	498
Biology	343	6.4%	227	6,222
Chemistry	169	7.7%	127	4,052
Computer Science	420	11.4%	335	18,193
Earth Sciences	127	5.5%	103	2,025
Ecology	243	4.1%	186	6,287
Economics	531	4.9%	428	16,109
Education	222	9.5%	182	7,565
Engineering	367	6.0%	284	19,713
History	23	8.7%	18	607
Language & Literature	73	5.5%	60	3,617
Law	56	5.4%	38	891
Library Science	31	6.5%	23	1,012
Mathematics	178	7.3%	138	5,665
Media & Communications	27	3.7%	22	396
Medicine	2,138	7.2%	1,589	56,596
Miscellany	269	7.8%	234	53,119
Other Sciences	320	10.9%	272	30,732
Philosophy	9	0.0%	4	79
Physics	104	2.9%	74	1,647
Political Science	62	3.2%	45	1,182
Psychology	54	1.9%	46	1,364
Religion	7	14.3%	5	185
Sociology	113	3.5%	94	3,512
Technology	305	9.2%	228	7,749
Zoology	154	2.6%	115	3,011

Table 8.1. Journals and 2015 articles by subject, alphabetic order

Subject	Journals	Articles	Art/Jrnl
Medicine	1,589	56,596	35.6
Miscellany	234	53,119	227.0
Other Sciences	272	30,732	113.0
Engineering	284	19,713	69.4
Computer Science	335	18,193	54.3
Economics	428	16,109	37.6
Agriculture	305	8,756	28.7
Technology	228	7,749	34.0
Education	182	7,565	41.6
Ecology	186	6,287	33.8
Biology	227	6,222	27.4
Mathematics	138	5,665	41.1
Chemistry	127	4,052	31.9
Language & Literature	60	3,617	60.3
Sociology	94	3,512	37.4
Zoology	115	3,011	26.2
Earth Sciences	103	2,025	19.7
Physics	74	1,647	22.3
Anthropology	47	1,614	34.3
Psychology	46	1,364	29.7
Political Science	45	1,182	26.3
Library Science	23	1,012	44.0
Law	38	891	23.4
History	18	607	33.7
Arts & Architecture	23	498	21.7
Media & Communications	22	396	18.0
Religion	5	185	37.0
Philosophy	4	79	19.8

Table 8.2. Subjects with the most articles, and articles per journal

Subject	Articles	GOAJ	Gray%
Miscellany	53,119	11,451	464%
Economics	16,109	15,859	102%
Other Sciences	30,732	33,488	92%
Engineering	19,713	28,044	70%
Computer Science	18,193	26,271	69%
Ecology	6,287	12,196	52%
Mathematics	5,665	11,239	50%
Technology	7,749	15,700	49%
Education	7,565	15,698	48%
Agriculture	8,756	21,939	40%
Library Science	1,012	2,874	35%
Medicine	56,596	173,922	33%
Language & Literature	3,617	13,298	27%
Zoology	3,011	11,137	27%
Chemistry	4,052	15,015	27%
Sociology	3,512	14,638	24%
Anthropology	1,614	6,926	23%
Political Science	1,182	5,196	23%
Psychology	1,364	6,457	21%
Earth Sciences	2,025	10,451	19%
Biology	6,222	33,140	19%
Law	891	5,239	17%
Arts & Architecture	498	5,075	10%
Media & Communications	396	4,560	9%
History	607	7,544	8%
Physics	1,647	22,493	7%
Religion	185	3,921	5%
Philosophy	79	3,336	2%

Table 8.3. Gray 2015 articles as percentage of *GOAJ* 2015 articles

A few notes on some of the interesting items in these tables:

- The highest *percentages* of inactive journals (ones that published during the 4.5-year period but not in 2015) are in Philosophy and Arts & Architecture—the only subjects with more than one-third inactive—but by far the largest *number* is in Medicine, which accounts for more than one-third of all inactive journals.
- Medicine has the most journals and articles: that's entirely typical of OA. Miscellany being so close behind, with 94% as many articles, is entirely *atypical*, since in *GOAJ* Medicine has fifteen times as many articles as Miscellany.
- The most prolific journals are in the two multidisciplinary subjects, not at all surprising since that's where most papermills are.
- Looking at total 2015 articles, Economics and Education are the only HSS fields other than Miscellany in the top 13 subjects—and the bottom ten are entirely HSS.

You can doubtless arrive at other interesting conclusions.

Segment by Segment

The next three chapters look at the three broad subject segments, offering a consistent set of tables and graphs for each one—a set that's largely comparable to that used in Chapters 9-11 of *GOAJ*.

9. Biology and Medicine

	Journals	Active2015	Articles	Art/Jrnl
Free	177	143	2,434	17
Pay	2,304	1,673	60,384	36
Total	2,481	1,816	62,818	35
Free%	7.1%	7.9%	3.9%	

Table 9.1. Journals and articles, biomed

What few no-APC (free) journals there are in biomed publish half as many articles per journal as the rest of them.

	Jan-Jn2016	2015	2014	2013	2012
Journals	1,637	1,816	1,497	1,130	663
%Free	6.8%	7.9%	8.6%	9.6%	9.0%
Articles	36,132	62,818	47,063	34,104	24,068
%Free	3.4%	3.9%	5.0%	6.1%	6.0%

Table 9.2. Journals and articles by year, biomed

Gray biomed grew rapidly from 2012 through 2015, and article count may still be growing—even as the free portion slips slowly away.

Peak Size	Journals	%Free	Articles	%Free
600+	16	0.0%	16,334	0.0%
150 to 599	69	2.9%	13,424	3.1%
60 to 149	179	2.2%	12,673	1.9%
20 to 59	506	11.5%	13,919	9.5%
1 to 19	1,046	7.6%	6,468	6.9%

Table 9.3. Article volume, biomed

Gray journals tend toward the very small (and very specific), but the big journals publish the most articles.

Charge	Jour.	%APC	%All	Art.	%APC	%All
$1,400+	166	9.9%	9.1%	11,173	18.5%	17.8%
$600-$1,399	464	27.7%	25.6%	9,645	16.0%	15.4%
$200-$599	678	40.5%	37.3%	10,512	17.4%	16.7%
$2-$199	365	21.8%	20.1%	29,054	48.1%	46.3%
Free	143		7.9%	2,434		3.9%

Table 9.4. APC levels, biomed

Nearly half of all fee-based articles appear in the lowest-journals, which also have the most articles per journal overall.

Figure 9.1, on the next page, shows starting date for gray biomed journals, most of which started quite recently.

Finally, Table 9.5 shows countries (clearly identifiable on publisher or journal sites) that published more than 150 articles in 2015 in gray OA journals. The obvious note: India alone accounts for roughly two-thirds of all the articles, even not including the thousands of "United States," "Canada" and "United Kingdom" articles that may come from India.

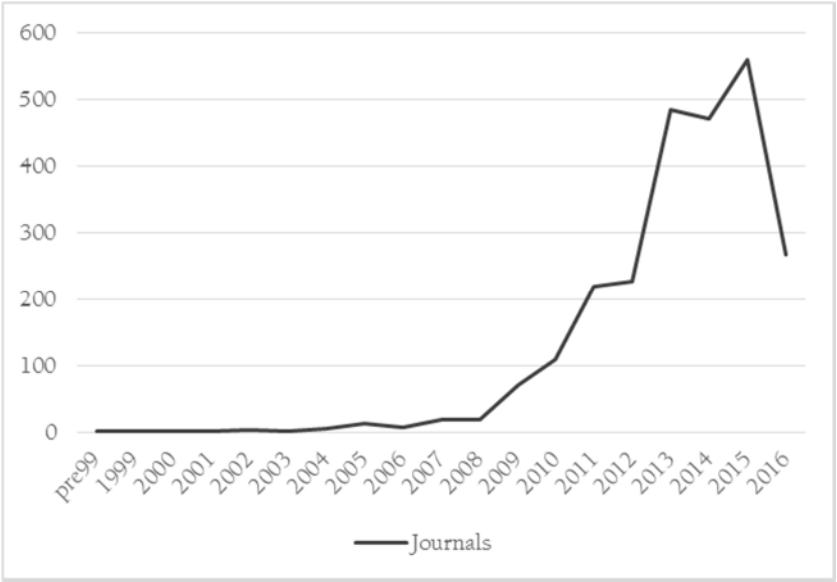

Figure 9.1. Starting dates, biomed

Country	Journals	Articles
India	820	41,460
United States	293	7,100
Canada	37	1,181
United Kingdom	46	1,027
Japan	1	896
Nigeria	119	788
Pakistan	61	733
Hong Kong	31	307
Australia	3	287
China	4	185
South Korea	1	170
United Arab Emirates	17	156
Turkey	9	154
Singapore	4	151

Table 9.5. Countries with more than 150 biomed articles in 2015

10. Science, Technology, Engineering and Math

The largest group of journals and articles—and with many more articles per journal than biomed.

	Journals	Active2015	Articles	Art/Jrnl
Free	194	157	3,160	20
Pay	2,586	2,010	104,670	52
Total	2,780	2,167	107,830	50
Free%	7.0%	7.2%	2.9%	

Table 10.1. Journals and articles, STEM

	Jan-Jn2016	2015	2014	2013	2012
Journals	1,876	2,167	1,961	1,688	1,059
%Free	8.0%	7.2%	7.3%	7.2%	6.7%
Articles	55,620	107,830	104,299	76,886	51,053
%Free	3.1%	2.9%	3.0%	3.3%	3.7%

Table 10.2. Journals and articles by year, STEM

After rapid growth from 2012 through 2014, growth has slowed but continues. There were never many free articles; that hasn't changed.

Peak Size	Journals	%Free	Articles	%Free
600+	33	0.0%	35,503	0.0%
150 to 599	172	2.9%	34,039	2.8%
60 to 149	279	2.9%	17,596	2.0%
20 to 59	646	8.2%	14,380	8.5%
1 to 19	1,037	8.8%	6,312	10.0%

Table 10.3. Article volume, STEM

Most journals are very small and most articles appear in large and very large journals.

Charge	Jour.	%APC	%All	Art.	%APC	%All
$1,400+	34	1.7%	1.6%	1,254	1.2%	1.2%
$600-$1,399	184	9.2%	8.5%	6,531	6.2%	6.1%
$200-$599	1,003	49.9%	46.3%	26,136	25.0%	24.2%
$2-$199	789	39.3%	36.4%	70,749	67.6%	65.6%
Free	157		7.2%	3,160		2.9%

Table 10.4. APC levels, STEM

More than two-thirds of fee-based articles involve very low fees, and, as with biomed, these are the journals with the most articles per journal.

Figure 10.1 on the next page shows the starting date of STEM journals; notably, new publications have slowed considerably since 2013.

Finally, Table 10.5 shows countries with more than 400 gray STEM articles in 2015; once again, India dominates the area with more than two-thirds of all articles.

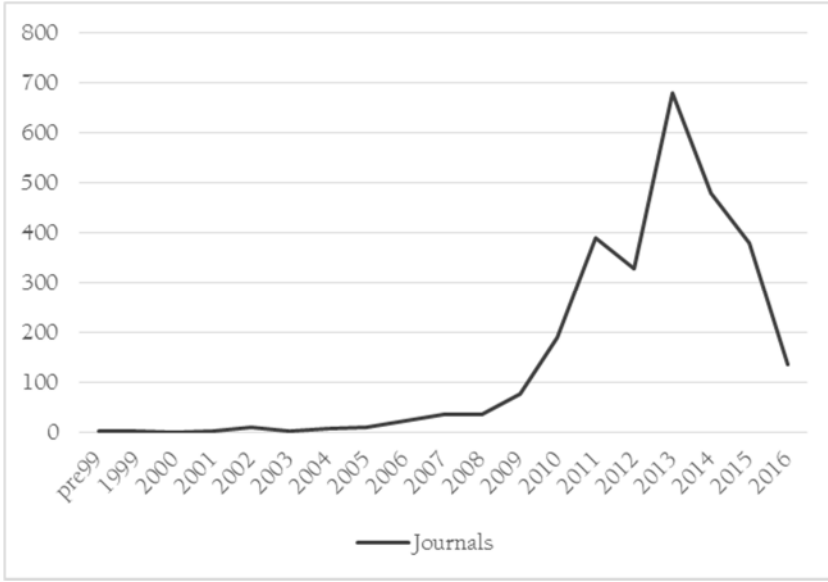

Figure 10.1. Starting dates, STEM

Country	Journals	Articles
India	825	72,604
United States	341	4,672
Canada	66	2,223
Pakistan	58	1,947
United Kingdom	83	1,642
Bulgaria	18	1,558
Nigeria	186	1,112
Bangladesh	10	949
South Korea	11	831
Morocco	2	791
Switzerland	10	539
Austria	3	536
United Arab Emirates	16	517
Singapore	10	496

Table 10.5. Countries with 400+ 2015 gray articles, STEM

11. Humanities and Social Sciences

The fewest active journals, an in-between number of articles—and by far the *most* articles per journal overall (but most of the articles are in the Miscellany group, including papermills).

	Journals	Active2015	Articles	Art/Jrnl
Free	96	84	3,118	37
Pay	1,484	1,185	88,632	75
Total	1,580	1,269	91,750	72
Free%	6.1%	6.6%	3.4%	

Table 11.1. Journals and articles, HSS

	2016	2015	2014	2013	2012
Journals	1,118	1,269	1,102	885	504
%Free	6.5%	6.6%	5.3%	4.9%	5.6%
Articles	47,575	91,750	72,601	51,942	32,091
%Free	2.5%	3.4%	3.3%	3.7%	3.3%

Table 11.2. Journals and articles by year, HSS

HSS gray OA grew rapidly in 2013 and 2014, with slower journal growth but substantial article growth in 2015.

Table 11.3 shows the familiar pattern of a few very large journals publishing a huge chunk of all articles (42% in this case)—but a lower percentage of very small journals.

Peak Size	Journals	%Free	Articles	%Free
600+	35	0.0%	38,335	0.0%
150 to 599	128	5.5%	27,054	5.2%
60 to 149	192	6.3%	12,934	5.3%
20 to 59	367	7.6%	9,588	7.6%
1 to 19	547	6.8%	3,839	7.3%

Table 11.3. Article volume, HSS

Charge	Jour.	%APC	%All	Art.	%APC	%All
$1,400+	3	0.3%	0.2%	132	0.1%	0.1%
$600-$1,399	48	4.1%	3.8%	2,291	2.6%	2.5%
$200-$599	519	43.8%	40.9%	28,769	32.5%	31.4%
$2-$199	615	51.9%	48.5%	57,440	64.8%	62.6%
Free	84		6.6%	3,118		3.4%

Table 11.4. APC levels, HSS

Almost no HSS gray OA journals charge high APCs, and as usual the very inexpensive journals publish the most articles per journal, more than twice as many as the two most expensive categories.

Figure 11.1 on the net page shows starting dates, and the pattern is somewhat similar to STEM but with less of a slowdown in 2014 and 2015. Once again, 2013 is the peak year for journal startups.

Finally, Table 11.5 shows stated country of publication for countries with at least 400 gray OA articles in 2015. While India still accounts for a majority of all articles, it's by far the smallest majority: 52%, as compared to 67% for STEM and 66% for biomed.

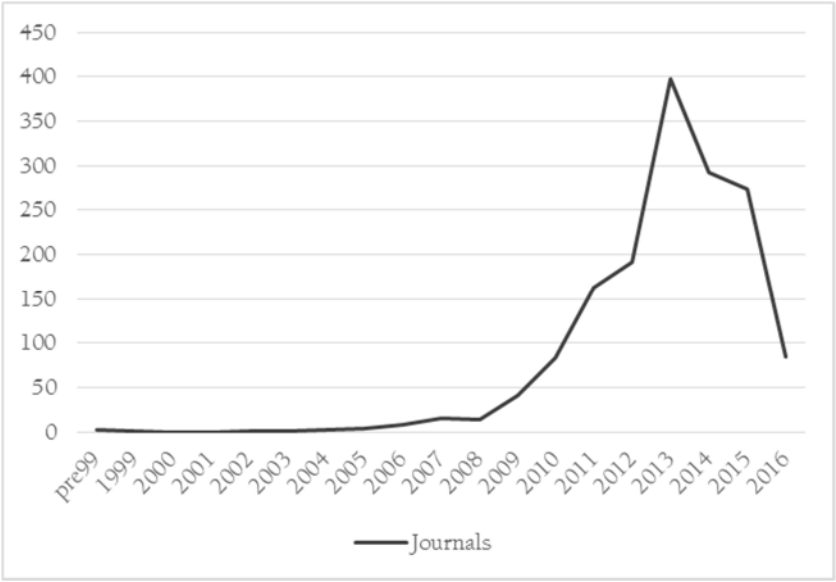

Figure 11.1. Starting dates, HSS

Country	Journals	Articles
India	387	47,810
United States	219	5,670
Canada	50	4,488
United Kingdom	72	3,131
Bangladesh	11	1,663
Romania	23	1,537
Pakistan	31	1,245
Nigeria	98	879
Turkey	23	879
Singapore	6	742
Australia	15	649
Austria	6	533
British Virgin Islands	1	521
Switzerland	3	414
Morocco	1	400

Table 11.5. Countries with at least 400 2015 articles, HSS

12. A Thought Experiment

To what extent is gray OA not only a regional phenomenon but a national phenomenon? Given that 62% of 2015 A and B articles with obvious country of publication were published in India, it's fair to say it's an issue.

What happens if you remove Indian publishers and journals, as well as those with questionable country assignments (Canada, the UK and the US) and those without obvious country statements?

That's what this chapter is about. I'm not impugning Indian publishers in general; after all, it has nearly 1.3 billion people, the world's third largest higher education system, and hundreds of thousands of researchers. But it's interesting to see what's left. Let's call this the "rest" group—the rest of the world where countries are identified.

Publishers

Eliminating the four countries mentioned, publishers where the country isn't obvious, and "publishers" (those with entirely empty, non-OA, DOAJ, or APC-hidden journals) leaves 150 publishers.

Those publishers produced a total of 1,314 A and B journals and 23 UA (APC hidden) journals, along with 1,224 empty "journals," 52 journals with X codings, and 259 journals in DOAJ.

The journals appeared to publish about 11,770 articles in the first half of 2016, about 24,790 in 2015, about 25,586 in 2014, about 21,508 in 2013 and about 17,257 in 2012.

Journals

Let's look at the A and B journals—1,403 in all, with 1,060 active in 2015—using the same tables as in Chapters 9-11 and adding one more table summarizing revenue and cost per article by segment.

	Journals	Active2015	Articles	Art/Jrnl
Free	170	134	3,321	25
Pay	1,233	926	26,440	29
Total	1,403	1,060	29,761	28
Free%	12.1%	12.6%	11.2%	

Table 12.1. Journals and articles, rest

On one hand, these are mostly small journals. While journals active in 2015 are 11% as numerous as in *GOAJ*, there are only 5% as many articles. On the other, unlike *GOAJ* or the overall gray OA picture, free journals have nearly as many articles per journal as APC-charging journals and published 11% of the articles, much better than the 3% for gray OA as a whole.

	2016	2015	2014	2013	2012
Journals	790	1,060	968	764	414
%Free	14.8%	12.6%	11.1%	10.3%	9.4%
Articles	14,010	29,761	29,533	25,766	19,195
%Free	13.6%	11.2%	9.8%	7.6%	6.7%

Table 12.2. Journals and articles by year, rest

The free percentage seems to be growing.

Peak Size	Journals	%Free	Articles	%Free
600+	9	0.0%	4,992	0.0%
150 to 599	61	11.5%	10,883	11.9%
60 to 149	93	8.6%	5,399	8.8%
20 to 59	228	19.7%	4,901	23.2%
1 to 19	669	11.1%	3,586	11.3%

Table 12.3. Article volume, rest

The few very large journals do *not* publish an outsize proportion of all papers.

Charge	Jour.	%APC	%All	Art.	%APC	%All
$1,400+	1	0.1%	0.1%	896	3.4%	3.0%
$600-$1,399	47	5.1%	4.4%	1,469	5.6%	4.9%
$200-$599	630	68.0%	59.4%	12,554	47.5%	42.2%
$2-$199	248	26.8%	23.4%	11,521	43.6%	38.7%
Free	134		12.6%	3,321		11.2%

Table 12.4. APC levels, rest

There's only one expensive journal (albeit a large one), and nine of ten fee-based articles appear in moderate-priced or low-cost journals.

Segment	Articles	Revenue	$/article
Biomed	4,782	$ 2,994,795	$ 626.26
STEM	12,649	$ 3,100,719	$ 245.14
HSS	12,330	$ 1,990,175	$ 161.41
Total	29,761	$ 8,085,689	$ 271.69

Table 12.5. 2015 revenue and cost per article by segment

Biomed is the smallest segment for these journals—but still brings in the most potential revenue, since the average cost per article is 2.5 times as much as for STEM and 3.9 times as much as for HSS.

The starting year graph for these journals is so similar to the STEM graph that it hardly seems worth printing.

Table 12.6, which ends this chapter, shows countries with at least 200 gray OA articles in 2015. It's not really new information.

Country	Journals	Articles
Pakistan	150	3,925
Nigeria	403	2,779
Bangladesh	22	2,613
Bulgaria	28	1,824
Romania	29	1,729
Singapore	20	1,389
Turkey	47	1,307
Australia	34	1,255
Morocco	3	1,191
Austria	10	1,152
Japan	5	1,131
South Korea	14	1,014
Switzerland	14	963
Malaysia	44	746
Iran	15	727
United Arab Emirates	36	711
Hong Kong	58	614
British Virgin Islands	1	521
Russia	2	476
Croatia	5	374
China	8	266
Bosnia and Herzegovina	4	254
Kenya	15	254
Poland	1	220
Tunisia	10	217
Germany	12	216
Italy	2	207

Table 12.6. 2015 journals and articles for remaining countries with 200+ articles

13. Comments and Conclusions

This is where I should comment on some the oddities among gray OA publishers and journals and offer sweeping conclusions.

I could certainly discuss the "clever" software that some gray publishers use to prevent one from using right-click options to open journal sites and archive segments in new browser tabs—usually with a silly warning about the site being copyright (or copy right). I could discuss journal sites that offer nothing that will make it easy to count articles—e.g., "PDF" or a DOI prefix or anything consistently findable. (There are others, fortunately, that simply number articles in each issue—or, better yet, number DOIs sequentially from the beginning of each year to the end or from the beginning of a *journal* to the end.)

I could spend more space on template publishers, and especially on the canned just-swap-in-the-core-of-the-journal-name "about" pages they use, which mostly don't make sense. (I continue to wonder whether "publishers" should lack the final "s," as most of these use *such* similar templates that I suspect one group is responsible for several of them. I'd call them the Merry Pranksters of Gray OA—except that a few authors have submitted papers to a few of the journals.)

Are these predatory journals? Not really, except for UA cases. (If journals state one APC and then ask for more money, that's flat-out fraud, a different issue.)

Are they questionable? At one level, yes: they're not in *DOAJ*. For many active journals, including the UA cases, I believe the answer is yes, but we've discussed that in Chapter 3.

But that's too simple. Questionable publishers may have good journals (just as some quality subscription journals have or have had highly questionable journals), and questionable journals may consist of mostly legitimate articles. As I looked at occasional articles and many article titles (noting that I'm unqualified to judge articles in most fields), my

sense was that most were narrow scholarship: not defective research but research that might not be welcomed in other journals.

International Journal…

Of the 18,910 journal titles I recorded (for various reasons, a few titles didn't make it into the spreadsheets), 3,277 began with "International Journal." That's 17%. But it gets worse: 25% of the A- and B-coded journals begin "International Journal." By comparison, only 6% of the journals in *GOAJ* begin "International Journal"—701 of 10,945.

I expected empty gray journals to have a high percentage of "International Journal"s—but only 10% of them begin that way. Turns out that's because most template publishers use different prefixes—so, for example, there are 520 empty journals with titles beginning "Academic Open," 351 beginning "American Open," 472 beginning "Asian American," 681 beginning "British Open," 430 beginning "Canadian Open," 529 beginning "Eurasian," 534 beginning "European Open," 527 "North American Open," 1,215 "Universal Open" and 450 "US Open." (See a pattern here?)

Coping with the Gray

India clearly has issues with scholarly publishing, given the sheer dominance of gray OA. Those issues may have to do with requirements for advancement in higher education or with a lack of awareness of the virtues of serious OA (where "serious" implies meeting *DOAJ* standards and becoming part of that directory).

Take away India and the "probably not" countries, and there's not much gray left, as discussed in Chapter 12. I suspect a lack of awareness is an issue in some countries, and I suspect that *DOAJ*'s new regional and national ambassadors will help rectify this situation. What can India, Nigeria and others learn from South America?

Ideally, the field of active gray OA would shrink to the point where it consists of truly questionable or even predatory publishers—but ideals are sometimes hard to achieve.

I would not fault any researcher for avoiding all gray OA journals: that's what I'd do if I was publishing research articles. But I can't entirely

fault the thousands of authors who've already appeared in gray OA journals. Suggestions that they're getting ripped off for huge sums are a bit overblown: after all, only 515 journals charged more than $1,000 at their highest rates, while most articles appeared in journals with fairly modest fees.

I suppose a word of thanks should go to Cenyu Shen and Bo-Christer Björk. If the numbers they proposed for "predatory" articles hadn't struck me as absurdly high, I wouldn't have spent several hundred hours doing a full survey of gray journals. But they did, I did, and here are the results. I hope you'll find them useful.

The Dataset

A portion of the master spreadsheet for this project will appear on figshare—not including subjects, country codes and some other material but including the counts, codes and APCs as I found them.

The dataset is available at https://figshare.com/articles/Gray_OA_2012-2016_Gold_OA_Beyond_DOAJ/4275860

Pay What You Wish

Cites & Insights carries no advertising and has no sponsorship. It does have costs, both direct and indirect. If you find it valuable or interesting, you are invited to contribute toward its ongoing operation. The Paypal donation button (for which you can use Paypal or a credit card) is on the *Cites & Insights* home page. Thanks.

Masthead

Cites & Insights: Crawford at Large, Volume 17, Number 1, Whole # 200, ISSN 1534-0937, a periodical of libraries, policy, technology and media, is written and produced irregularly by Walt Crawford.

Comments should be sent to waltcrawford@gmail.com.

URL: citesandinsights.info/civ17i1.pdf