Gray OA 2014-2017:
A Partial Follow-Up

Walt Crawford

A Special Issue of

Cites & Insights

Crawford at Large

Libraries • Policy • Technology • Media

Volume 17, Number 9: October 2017 ISSN 1534-0937

Contents

1. Methods and Changes

This informal study extends *portions* of *Gray OA 2012-2016: Open Access Journals Beyond DOAJ* (published as *Cites & Insights* 17.1), but neither replaces that study nor attempts to provide a full update. If you haven't read the earlier study, you should; this follow-up will make better sense if you do. Chapters 2-5 follow the general approach of Chapters 1-3 and 5 of the earlier study, and Chapter 8 follows the general approach of Chapter 13. Chapters 6 and 7 use the same subjects and countries as Chapters 7 and 8 of the earlier study, but for different and much narrower purposes.

This time, instead of offering the big picture right up front, I'll say a little about how this follow-up was prepared and the resulting changes.

The Starting Point and Data Gathering

I began with the master dataset used for *Gray OA 2012-2016*, minus 50 journals that proved to be duplicate entries (either different languages or different titles with the same URL, leaving 18,860 journals and "journals." I did *not* add new non-*DOAJ* journals appearing since July 1, 2016 or ponder any additions to the "ppppredatory" lists.

This time around, I ignored questions of fee or free and the amount of APC if any. It was already clear that, unlike *DOAJ*, the vast majority of active gray OA journals *do* charge (usually small) APCs—considerably more than 90%—and I already knew that tracking down APCs is time-consuming. So, with one exception, this follow-up concerns article volume and site availability, *not* APC-or-free status or APC amount. The exception: in cases where a journal had a hidden or unknown APC, coded "UA" in the earlier report but since changed to "FA" to conform to *GOAJ2*, I checked to see whether a fee was clearly visible. If it was, I changed the coding appropriately.

The process this time around:

- Starting with all journals sorted first by publisher, then by journal. I attempted to visit each URL. Exceptions: if the first three journals from

a multijournal publisher all yielded code XX (unavailable) or XM (malware), and if the URLs were of the typical publisher/title form (that is, a root domain for the publisher followed by a page for the journal) I checked the publisher site itself. If it showed the same condition, I replicated the code across the remaining journals without checking each one. This probably affected fewer than 50 journals in all.

- For those journals that were available, I took article counts for all of 2016 and the first half of 2017. In cases where a journal "recovered" (going from an excluded category such as XX or XM to a workable status), I also counted articles for 2015 and, in some cases, 2014. Counts for 2017 were doubled: note that the column heading is "2017x2." Based on the difference between 2016 estimates in the first report and actual 2016 counts, it's likely that actual 2017 counts will be within 5%-10% of the doubled first-half numbers *in general*, although specific journals will show wide variances.
- I used Edge rather than Chrome this time around, for two reasons—the second far more important. First, Edge *seems* to be faster and less resource-intensive (and doesn't just stall for 30 seconds every so often). (The primary reason to use Chrome rather than Firefox, built-in page translation, is irrelevant: if anything, Edge's translate-this-page facility is a bit more convenient and works well enough.) Second, the way Edge does ctrl-F (find on this page) is *much* more efficient when you need to do the same Find over many pages (e.g., "PDF" on each of 12 issues for each of ten journals).
- Working publisher-by-publisher made it *much* easier to count articles, since most publishers use similar metadata and formatting for all journals.
- Unlike the earlier report, where if a publisher had 300 or more "journals" and the first 100 had no articles at all I would mark them all as empty, I visited every journal this time around. However, for publishers that clearly didn't add descriptive paragraphs to the home page until the first article was published—the pattern for several "publishers" that may all actually be a single entity, since they use essentially identical templates—I did not actually go to "current issue" for all the essentially empty home pages.
- In a few cases, I had to use an approximation or gave up on doing a count. There were fewer of those cases than in the previous report.

- I did *not* pick up country codes or assign subject codes for journals that didn't already have them (for subjects, those journals that weren't available or were empty the last time around). That only matters for the discussions of volatility.
- Most counts and tables *do* include "FA" journals, those with hidden or unstated APCs. While I regard such journals as inherently questionable (perhaps the clearest example of "predatory"), since this update is about article volume rather than revenues it makes sense to include them.

Expectations and Reality

Given the increasing number of alarmist articles about the dangers of ppppredatory publishing, I frankly expected to see a substantial drop in article numbers and active journals.

That didn't happen—or at least it's not that simple. There may be several reasons, for example:

- Many of the "ppppredatory" journals aren't predatory at all but have not seen fit to apply for *DOAJ* listing; they have satisfied groups of authors who continue to publish and encourage others to do so.
- Some scholars have never heard the whole heated discussion around "ppppredatory" journals or have concluded that they don't care about it.
- Some scholars are using these journals as easy ways to publish.
- I'm fairly certain that India's UGC list has made a difference: I saw the "UGC" label in a fair number of journals with sharply-*increasing* article counts. Someone with more stamina or a way to download the massive UGC list might investigate this further; in Chapter 6, I use India as a UGC surrogate, which is simplistic.

The reality? Gray OA, at least for this huge subset, is reasonably stable: possibly shrinking *slightly* in 2017 (less than 1%) after slow growth in 2016 (less than 6%).

Oh, and 70 journals that were gray OA in July 2016 are now in *DOAJ*—around 1% of the "real" journals and around 1% of the articles.

2. 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 *GOAJ2: Gold Open Access Journals 2011-2016* (henceforth *GOAJ2*), 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. The earlier gray OA report got 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.) This follow-up offers complete article counts for 2016 and figures for the first half of 2017, doubled to make comparisons easy.

Herewith, then, some oversimplified figures for gray OA, offered comparably to those on page 1 of *GOAJ2:*

- Gray OA journals that were fully available in the summer of 2017 published 296,122 articles in 2017 (extrapolated from half that number through June 30); 298,215 in 2016; 282,845 in 2015; and 235,370 in 2014. (The 2014-2016 numbers differ from those in the earlier report because of journals that were no longer available or that had previously been unavailable.)
- In all, 7,860 currently-available gray OA journals published at least one article between January 1, 2012 and June 30, 2017, so you *could* say there were an average of 38 articles per journal in 2017—but that's misleading.
- There are a staggering 18,790 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.

- Including only journals that actually published articles in a given year (or half-year doubled), 4,963 journals published 296,122 articles (extrapolated) in 2017. That's an average of 60 articles per journal. For 2016, a *lot* more journals (6,086) published just slightly more articles (298,215), for an average of 49 articles per journal. For 2015, the active journal count was slightly lower than for 2016 (5,653), as was the article count (282,845), for an average of 50 articles per journal. Finally, the too-low figures for 2014 (because, for some journals visible in summer 2017 that weren't visible in winter, I didn't take the counts back this far—the difference might amount to 2% or less), I show 4,748 journals and 235,370 articles, or an average of 50 articles per journal.
- For that matter, if I included articles counted earlier for journals that are no longer visible (malware, 404s, DSN failures, etc.), it would add 8,027 articles in 2016; 20,358 in 2015; and 21,809 in 2014. (Journals now in *DOAJ* account for 2,178 articles—extrapolated—in 2017; 3,397 in 2016; 3,652 in 2015; and 4,080 in 2014.)

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 in more detail. As appropriate, I'll include *GOAJ2* figures and grand totals for 2014-2016—noting that such totals still aren't quite comprehensive. Still: more than 823,000 articles in gold OA journals in 2016 (including the 3,397 in journals now in *DOAJ*): that's a striking number, nearly a third of the presumed 2.5 million total scholarly articles per year.

The Biggest Numbers

I *do* include journals with hidden/unstated APCs in most discussions because they represent a larger portion of the whole. In 2016, for example, such journals accounted for 31,704 articles (10.6% of the total), compared to just over 1% for *DOAJ*.

Table 1.1 shows article counts for journals counted in this report, with codes for a number of special cases.

Code	Journals	2017x2	2016	2015	2014
A	3,764	265,221	250,072	235,571	192,897
B3	261				
B4	392				2,015
B5	542			2,966	2,591
B6	1,367		11,969	12,584	10,497
BC	76		4	11	14
BF	768	2,198	3,694	4,900	4,525
BR	12	915	772	589	803
FA	678	27,788	31,704	26,224	22,028
Total	**7,860**	**296,122**	**298,215**	**282,845**	**235,370**

Table 2.1. Articles per year and codes

Notes on the codes:

- "A" is the catchall code for journals that didn't get another 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 are most likely defunct.
- "B5" journals haven't published any articles since 2015. Some of these are on hiatus; most are probably defunct.
- "B6" journals published articles in 2016 but not in the first half of 2017. As you may note in the earlier report for "B5," this code has many more journals than earlier codes—and there are fewer than half as many "B5" as there were in that report. Most likely, most of these have very long lead times.
- "BC" journals fall into one of two categories: explicitly ceased or merged into other journals (thus the 2014-2016 numbers), or with no articles more recent than 2012. It seems fair to assume that a journal with no activity in 4.5 years is defunct.
- "BF" journals have either one or two articles in the first half of 2017 (extrapolated to 2 or 4), too few to represent robust publishing.

- "BR" journals consist entirely or primarily of reviewed papers presented at conferences.
- "FA" journals fail for lack of visible APC amounts.

These codes are directly comparable to those used in *GOAJ* and *GOAJ2* (where "FA" is coded "CA"). There are proportionally more of most "B" codes: gray OA journals are more erratic in general.

Journal Stability

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? How many are "real journals" rather than one-year wonders or skeletal "journals"?

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 and, later, more than nine (using extrapolated counts for 2017).

Years with Five or More Articles

Years > 4	2017x2	2016	Cum%	2015	2014
None	314	625	10.3%	673	672
One	577	798	23.4%	623	520
Two	780	1,073	41.0%	764	479
Three	880	1,178	60.4%	1,181	665
Four	2,412	2,412		2,412	2,412
Total	**4,963**	**6,086**		**5,653**	**4,748**

Table 2.2. Gray journals publishing five or more articles per year

Table 2.2 shows the number of gray OA journals (coded A, B or F) 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 2017).

If you define two active years as minimal for a stable journal, most gray journals with articles in 2016 make it: about 59%. If four years is

the target, just under 40% manage. Note also that, out of 7,860 journals, there's never a year without at least 1,774 not publishing *any* articles. (I used 2016 because those are full-year figures; if 2017 extrapolations are accurate, the percentage of stable journals among those actually publishing is even higher.)

Years > 4	2017x2	2016	Cum%	2015	2014
None	870	1,275	0.4%	1,339	1,360
One	5,964	4,102	1.8%	3,383	4,398
Two	18,144	17,482	7.7%	7,761	5,554
Three	36,274	40,748	21.3%	26,429	12,961
Four	234,870	234,608		243,933	211,097
Total	**296,122**	**298,215**		**282,845**	**235,370**

Table 2.3. Articles in gray journals publishing at least five articles per year

Table 2.3 is in some ways more interesting than Table 2.2, as it demonstrates that stable journals tend to be larger journals overall. The 59% of active-in-2016 journals that published at least five articles in at least three years accounted for 92.3% of all 2016 articles, and those with four such years accounted for almost four out of five. (The percentages are even higher for 2015 and 2014.)

Years > 4	2017x2	2016	2015	2014
None	2.8	2.0	2.0	2.0
One	10.3	5.1	5.4	8.5
Two	23.3	16.3	10.2	11.6
Three	41.2	34.6	22.4	19.5
Four	97.4	97.3	101.1	87.5
Total	**59.7**	**49.0**	**50.0**	**49.6**

Table 2.4. Articles per journal for gray journals with at least five articles per year

Table 2.4 is derived from Tables 2.2 and 2.3, showing average articles per year (for active journals in each year) relative to number of years with at least five articles. You can draw your own conclusions (but see Chapter 5, since "average journal size" is a nonsensical phrase). Table 2.4, especially the "One," "Two" and "Three" rows, may suggest that the

final 2017 publication figures will be somewhat *lower* than the figures extrapolated here, but not massively so.

Raising the Bar and a DOAJ Comparison

Tables 2.5 through 2.7 are based on journals with at least ten articles in a given year. The patterns are similar, naturally. Total lines are omitted because they're inherently identical to Tables 2.2-2.4.

Years > 9	2017x2	2016	Cum%	2015	2014
None	858	1,525	25.1%	1,484	1,331
One	785	948	40.6%	769	546
Two	792	949	56.2%	736	459
Three	714	850	70.2%	850	598
Four	1,814	1,814		1,814	1,814

Table 2.5. Gray journals publishing ten or more articles per year

Years > 9	2017x2	2016	Cum%	2015	2014
None	3,660	5,430	1.8%	5,348	5,101
One	9,782	7,774	4.4%	6,575	6,475
Two	23,072	21,689	11.7%	10,013	7,638
Three	34,834	39,559	25.0%	28,500	14,768
Four	224,774	223,763		232,409	201,388

Table 2.6. Articles in gray OA journals publishing more than ten or more each year

Years > 4	2017x2	2016	2015	2014
None	4.3	3.6	3.6	3.8
One	12.5	8.2	8.6	11.9
Two	29.1	22.9	13.6	16.6
Three	48.8	46.5	33.5	24.7
Four	123.9	123.4	128.1	111.0

Table 2.7. Articles per journal for journals with at least ten articles per year

If you choose to sum this up as "there are just over 1,800 gray OA journals that consistently publish at least ten articles per year, and those journals average more than 120 articles per year, far more than less consistent journals," I think that's about right.

Comparative Figures for DOAJ

Tables 2.8-2.10 are directly comparable to Tables 2.5-2.7, except that they're for journals in *DOAJ* and consist of full-year figures for 2016, 2015, 2014 and 2013. For comparison purposes, the percentage figures are for 2016, as in Tables 2.5 and 2.6. *Only* articles in 2013-2016 were included for stability measures, ignoring 2011-2012 figures.

Years > 9	2016	Cum%	2015	2014	2013
None	369	4.4%	433	425	368
One	526	10.6%	494	470	395
Two	891	21.2%	976	761	588
Three	1,317	36.8%	1,450	1,452	966
Four	5,328		5,328	5,328	5,328
Total	**8,431**		**8,681**	**8,436**	**7,645**

Table 2.8. DOAJ journals publishing ten or more articles per year

The differences should be clear. Nearly two-thirds of the active journals published ten or more articles every year, compared to less than one-third of gray OA journals, and more than 19 of every 20 *DOAJ* journals published at least ten articles in one of the four years, compared to three out of four gray OA journals.

Years > 9	2016	Cum%	2015	2014	2013
None	1,909	0.4%	2,279	2,360	2,052
One	8,404	2.0%	4,119	4,031	3,988
Two	20,675	5.9%	18,094	10,236	8,464
Three	45,584	14.6%	49,700	42,774	17,886
Four	446,633		412,319	402,585	362,666
Total	**523,205**		**486,511**	**461,986**	**395,056**

Table 2.9. Articles in DOAJ journals publishing more than ten or more each year

Roughly six of every seven articles appeared in stable journals, compared to three out of four articles in gray OA journals.

Years > 9	2016	2015	2014	2013
None	5.2	5.3	5.6	5.6
One	16.0	8.3	8.6	10.1
Two	23.2	18.5	13.5	14.4
Three	34.6	34.3	29.5	18.5
Four	83.8	77.4	75.6	68.1
Total	**62.1**	**56.0**	**54.8**	**51.7**

Table 2.10. Articles per journal for DOAJ journals with at least ten articles per year

While less stable journals still publish fewer articles per journal, the differences for stable journals are less dramatic, probably because most *DOAJ* journals *are* stable.

Journal Growth and Shrinkage

Change 2015-16	Count	Percent	Cum%
Grew 50%+	2,291	32.1%	
Grew 15-49%	596	8.4%	40.5%
Even, ±14.99%	1,039	14.6%	55.1%
Shrank 15-49.9%	1,199	16.8%	71.9%
Shrank 50%+	2,006	28.1%	
Total	**7,131**		

Table 2.11. Gray OA journal growth and shrinkage 2015-2016

Table 2.11 shows growth and shrinkage for journals with articles in 2015, 2016 or both. Those that had articles only in 2016 are in the "grew 50%+" category and make up most of that category; those that had articles in 2015 but not in 2016 are in "shrank 50%+" and also make up most of that category.

The only real message here is that, even using a loose definition of "even," fewer than one of six journals stayed fairly even.

Change 2015-16	Count	Percent	Cum%
Grew 50%+	1,516	23.0%	
Grew 15-49%	531	8.1%	31.1%
Even, ±14.99%	1,011	15.4%	46.4%
Shrank 15-49.9%	1,115	16.9%	63.4%
Shrank 50%+	2,413	36.6%	
Total	**6,586**		

Table 2.12. Speculative gray OA growth and shrinkage 2016-2017

Table 2.12 gives similar figures for 2016-2017, using extrapolated figures for 2017; while most of the numbers probably aren't too far off, I'd be surprised if the "shrank 50%+" category didn't shrink as a few hundred (perhaps 300?) small journals with very long lead times publish their first 2017 articles in July-December 2017.

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 3 discusses the very large number of "journals" that aren't counted.

Chapter 4 peels the layers of the two source lists, specifically considering journals that aren't questionable OA at all. Chapter 4 also provides some comments on and measures of legitimately questionable journals.

Chapter 5 considers journals by article volume.

Chapters 6 and 7 look briefly at changes and volatility by country and subject, for the subset of journals where that information was captured.

The last chapter offers brief comments and conclusions.

3. Exclusions and Changes

Consider the journals with codes other than A, B and F—the ones excluded from most analysis. Articles in these journals weren't counted, in most cases because there was nothing to count.

This chapter breaks down the majority of gray "journals" that aren't countable (and the 678 that are counted but distinctly questionable). There are also some notes on changes in journal codes from the 2016 study to this one—most but not all of the changes positive.

The Codes—and a *GOAJ2* Comparison

Code	Journals	% of Norm
FA: Unknown or hidden APC	678	16.7%
XE: Empty since at least 2012	8,538	209.7%
XH: Hybrid	81	2.0%
XM: Malware	242	5.9%
XN: Not open access	131	3.2%
XO: Opaque, too difficult to count	47	1.2%
XU: Unworkable site	44	1.1%
XX: Unreachable or parking/ad	1,847	45.4%
Total excluded (w/o FA)	**10,930**	**268.4%**

Table 3.1. Journal exclusions for gray OA

Table 3.1 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 4,072 that have published five or more articles in at least two years and published at least one article in 2016.

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 *GOAJ2*: nearly eight times as many journals and roughly fifteen times the percentage of normal journals, 58.7% compared to 3.9%.

As an indication of just *how* startling the percentages are, Table 3.2 uses Table 3.1 from *GOAJ2*, modified to match the categories in Table 3.2 here (splitting XH out of XN, adding XI, XT and XV into XO, and adding XP into XX) and adds a % of Norm column based on the norm for *GOAJ2*: 7,776 journals. Table 3.2 should be fully comparable to Table 3.1. (Note that there were *no* translation failures in the gray OA analysis—partly because very few gray OA journals have non-English interfaces, although hundreds or thousands have interfaces with remarkably poor English.)

Code	Journals	% of Norm
FA: Unknown or hidden APC	40	0.5%
XE: Empty since at least 2011	46	0.6%
XH: Hybrid	4	0.1%
XM: Malware	67	0.9%
XN: Not open access	13	0.2%
XO: Opaque, too difficult to count	51	0.7%
XU: Unworkable site	21	0.3%
XX: Unreachable or parking/ad	147	1.9%
Total excluded (w/o FA)	**349**	**4.5%**

Table 3.2. Journal exclusions for *GOAJ2*

FA: 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.

Fifteen publishers had ten or more FA journals in 2017, accounting for two-thirds of all FA journals: JSciMed Central; SM Group Open Access Journals; ClinMed International Library; Universal Research Publications; SciDoc Publishers; Medwell Journals; SciRes Literature; Modern Scientific Press; Science Alert; Lawarence Press; PaperSciences Research Publisher; Verizona Publisher (VZP); Apex Journal; Merit Research Journals; and AENSI (American-Eurasian Network for Scientific Information). It's possible that some of these state fees at the publisher level, but it's never obvious at the journal level—and in far too many cases there's a statement about "nominal" fees that boils down to "trust is." (Here as elsewhere in these discussions, publishers are arranged by descending order of number of journals with these codes.)

XE: Empty from 2012 through June 2017

Most of this enormous group is "journals," entities that have *never* published any articles, although there are some that once had articles but faded away before 2012 and a few that may start publishing in the second half of 2017.

Fourteen publishers (several of them possibly the same entity, using essentially identical templates for their "journals") account for nearly 83% of all empty journals: 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); Research and Knowledge Publication; Science and Technology Publishing; and Journal Network. Each of these lists more than 290 empty "journals."

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.

Apart from four singleton journals, there are only seven publishers involved, and the first accounts for nearly two-thirds of the journals: OMICS International; Brainy Buzz; KEI Journals; iMed.pub; Business

Perspectives; Lawarence Press; Global Society of Scientific Research and Researchers (GSSRR).

I should note that some (perhaps most) of these journals are up front about being hybrid. They simply don't belong in a study of fully OA journals; they're here because of defective source lists.

XM: Malware

This is an astonishingly high number, especially since some journals that would have been flagged as malware when I researched *GOAJ2* wound up as unreachable instead. (McAfee SiteAdvisor doesn't seem to run in Edge, and had a tendency to flag some unreachable URLs as malware.) So, if anything, the comparison with *GOAJ2* is worse than it looks. Do note that, if a publisher's site used as a root for journal URLs was itself infected with malware, I flagged all of its journals as malware after spot-checking two or three.

Eight publishers with eight or more journals or root URLs flagged as malware account for 81% of the total cases: Global Science Research Journals; TLEP Journals (The Leading Edge Journal Publication Company); Ommega Publishers; Centre For Info Bio Technology (CIBTech); Revotech Press; International Recognition Multidisciplinary Research Journals, Monthly Publish; Pharma Research Library; and Wyno Academic Journals. (Seventeen singleton journals suffer from malware.)

XN: Not Open Access

Journals were flagged as not OA either because they label themselves as subscription, have embargos or require registration—or because attempts to open articles were met with refusals of some sort or an inability to get from abstracts to full text.

Nine publishers have five or more journals that do not appear to be OA peer-reviewed journals, accounting for slightly less than half of the total: Association for Sustainable Education, Research and Science (ASERS); Blue Eyes Intelligence Engineering & Sciences Publication; Lawarence Press; ABC Journals; Bowen Publishing; International Journals of Multidisciplinary Research Academy; Watch Plus; ASD Publisher; and eu-print.

XO: Opaque, too difficult to count

Note that most XO entries in Table 2.2 (*GOAJ2*) are actually merged journals that can no longer be counted individually; that's not an issue for gray OA, as far as I can tell.

Only two publishers had more than two journals I found it too difficult to count, accounting for roughly two-thirds of the total: Institute of Research Engineers and Doctors (IRED) and Convergence Information Society.

XU and XX: Unworkable or Unreachable

I now believe the distinction between these two is arbitrary; think of them as totaling 1,891 journals that couldn't be reached or just didn't work—as compared to 169 for XU and XX combined in *GOAJ2*. If a publisher was unreachable (after checking three or four journals that couldn't be reached and used the same root URL), I flagged all of its journals as unreachable.

Unfortunately, while most excluded categories (except malware) have improved since the 2016 study, the number of unreachable journals has more than doubled.

The most common clear reasons for XX include 404 errors (885 of them!); ad pages, parking pages and suspended accounts (282 in all), DNS errors or other similar failures (482) and not being findable from a parent page (96). There were other problems such as pages with no contents, looping menus, PDFs that never finished loading and journals that have morphed into entirely different things.

Nineteen publishers had at least 20 XX journals each, accounting for almost exactly half of the total: APST Publication; Advanced Research Publications; Wudpecker Research Journals; International Digital Organization for Scientific Information (IDOSI); Scientific Journals International; Swift Journals; Literati Scientific and Publishers (Literati Publishers); Insight Knowledge; World Academic Research Journals (WARJ); Eko Journal; International Association for Engineering and Management Education (IAEME); Center of Advanced Scientific Research and Publications (CASRP); German Science and Technology Press; American V-King Scientific Publishing; Horizon Journals; International Scholars Journals; Madridge Publishers; Cloud Journals; Enliven Archive; and Council for Innovative Research. Some of these have disappeared entirely.

Changes from July 2016 to July 2017

16||17	A	B3	B4	B5	B6	BC	BF	BR	FA	XE	XH	XM	XN	XO	XU	XX
A			1	7	445		263	2	1		1	60	18	2	7	193
B2							1									
B3	2		3	3	11	1	4		1			12	3			48
B4	14			15	22		25		3			21	1		1	68
B5	141		4		243	2	119		3		1	58	8	1	3	137
BC	1	1			3		3									29
BF	161			5	344	2			1			24	2		3	63
BR	11			1	7											
FA	88		5	23	83		10	5		4	1	10	3		3	86
XE	380	4	16	34	177	4	126		89		2	29	10	2	11	638
XH	2															41
XM	11			2	8		1		3				1			16
XN	8		1	1	2						6	1				32
XO	1											3	1			24
XU	2											1				15
XX	199	1	2	3	22	1	21		4	1		11	3	1	11	

Table 3.3. Code changes from 2016 to 2017

Table 3.3 shows code changes for journals from 2016 (rows) to 2017 (columns). The table may be difficult to read or interpret. A few notes:

- Good news: all the journals that are now in the "A" column, especially those that were previously "XM" or "XX.".
- Bad news: Anything in the XM column.
- Unsurprising: More XX journals, especially 638 XE "journals."

The B5-to-B6 number represents a couple hundred journals that appear to publish at least a year late.

.

4. Breaking Down the Lists and Questionable Journals

As with *Gray OA 2012-2016*, the now-discontinued Beall lists of publishers and journals formed the universe for this study (omitting journals in *DOAJ*)—not because such listing actually means that journals or publishers are questionable, but because it offered a universe to explore. Chapter 3 of the earlier work looked at publishers that didn't even make it into the full study (395 of them). That work is not repeated here.

While Beall failed to offer any evidence whatsoever for including most (approximately eight out of nine) of the publishers and journals on the lists, he did offer *some* evidence for *some* of them. Some others have clear evidence of questionable attitudes, and of course there are the truly predatory "FA" cases.

In one sense, every journal in this study (except those founded in 2016) is somewhat questionable, the 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 "FA" 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 evidence of publishing random articles with absurdly short review periods.

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.

Gray OA 2012-2016 had another category, "S," for journals in which all or nearly all articles were by the same single author. Only one of these journals published any articles after 2014, and that exception was by a different author, so I've omitted the category.

An important caveat here: Good papers appear in questionable journals, especially when the question is "how much is the APC?" 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.

Category	**2017x2**	**2016**	**2015**	**2014**
A: APC missing/hidden	329	463	456	360
B: Beall evidence	1,731	2,102	1,815	1,605
C: Crackpottery	6	7	6	6
L: Lorem ipsum		9	8	11
P: Papermill	77	78	79	67
Questionable total	**2,143**	**2,659**	**2,364**	**2,049**
Others	2,820	3,427	3,289	2,699
Total	4,963	6,086	5,653	4,748
Questionable %	43.2%	43.7%	41.8%	43.2%

Table 4.1. Questionable journals

Category	**2017x2**	**2016**	**2015**	**2014**
A: APC missing/hidden	20,256	23,454	21,639	18,181
B: Beall evidence	57,886	65,342	72,720	62,851
C: Crackpottery	384	365	463	509
L: Lorem ipsum		20	29	40
P: Papermill	61,452	59,685	52,011	32,229
Questionable total	**139,978**	**148,866**	**146,862**	**113,810**
Others	156,144	149,349	135,983	121,560
Total	296,122	298,215	282,845	235,370
Questionable %	47.3%	49.9%	51.9%	48.4%

Table 4.2. Articles in questionable journals

Tables 3.3 and 3.4 summarize the situation, and they're fairly revealing. Among other things, it's worth noting that—while legitimately questionable journals publish around half of gray OA articles—cases where Beall made a legitimate case accounted for than half of questionable cases and less than one-quarter of all gray OA articles in 2016. Also noteworthy: there aren't a lot of papermill journals but they churn out a lot of articles, and that number seems to be increasing. Finally, the two smaller questionable categories are so small they might not be worth mentioning, never totaling even 600 articles in a year.

5. Article Volume

This is the last chapter to include journals with missing or hidden APCs. It's also the last chapter looking at the whole analyzed universe; the next two chapters discuss the subset that already had country or subject codes from the previous report (and is limited to journals coded A or B).

Most gray journals don't publish very many articles, although there are exceptions. Six journals (up from two last year!) published more than 4,000 articles in their peak year (2014-2017), six more published more than 3,000 (up from one) and 10 more broke the 2,000-article mark (down from 11)—but only two journals published 2,000 articles in each of four years (up from zero, and it's worth noting that one of those two had malware last year but not this).

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 *GOAJ2*. Note that journals with *no* articles 2014-2017 (that is, code B3 and most of code BC) have been eliminated. That doesn't change article counts or annual active-journal counts; it does reduce total journals to 7,530.

Perhaps it's best to point out the big change here from *Gray OA 2014-2017*: the rise of the megajournals, most but not all "interdisciplinary" and about half with signs of being papermills. In 2014, journals with at least 2,000 articles published a total of 13,895 articles. For 2015, that figure more than doubled 33,806. For 2016, the total was 36,366—and based on article counts for the first half of 2017, the total is likely to be around 53,410. For that matter, given that some of these journals are growing so rapidly, it's likely that the final 2017 numbers will be even higher.

Roughly Equal Journal Numbers

Table 5.1 breaks down gray journals (excluding X codes) into ten roughly equal parts based on peak number of articles—"roughly" because 753-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 four 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 (100 or more articles), and even there at least seven of 755 journals were wholly absent in any given year. The worst case is the lowest category, 828 journals that never published more than two articles per year: no more than 369 of these, 44%, published in any given year.

It's also important to note that journals and articles are counted for each year in the group based on peak articles. To give an extreme example of how this affects the results, consider the top row of tables 5.3 and 5.4. They do *not* say that five journals published 4,000 or more articles in 2014: they say that five journals *with a peak article count of 4,000 or more* published at least one article in 2014. (Otherwise, the article counts for 2016-2014 would be impossible!)

Articles	Journals	2017	2016	2015	2014	%	Cum%
100+	755	732	748	727	653	10.0%	10.0%
51-99	743	703	726	690	618	9.9%	19.9%
32-50	766	688	731	679	574	10.2%	30.1%
22-31	748	628	704	626	509	9.9%	40.0%
15-21	758	589	684	598	449	10.1%	50.1%
12-14	697	475	600	535	414	9.3%	59.3%
8-11	898	538	701	606	453	11.9%	71.2%
5-7	772	296	567	519	406	10.3%	81.5%
3-4	565	185	343	331	303	7.5%	89.0%
1-2	828	129	282	342	369	11.0%	44.6%
Total	7,530	4,963	6,086	5,653	4,748		

Table 5.1. Journals grouped by roughly equal peak-article size.

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 "0.0" and every number in the **%** column would be 10.0.

Articles	2017	2016	2015	2014	%16	Cum%
100+	216,958	208,493	197,393	161,123	69.9%	69.9%
51-99	32,218	36,283	35,690	30,142	12.2%	82.1%
32-50	18,448	19,762	18,986	16,765	6.6%	88.7%
22-31	10,936	12,295	10,798	10,106	4.1%	92.8%
15-21	7,414	8,179	7,439	6,292	2.7%	95.6%
12-14	4,254	5,270	4,952	4,186	1.8%	97.3%
8-11	3,744	4,367	3,967	3,357	1.5%	98.8%
5-7	1,280	2,291	2,281	2,039	0.8%	99.6%
3-4	612	891	880	874	0.3%	99.9%
1-2	258	384	459	486	0.1%	
Total	296,122	298,215	282,845	235,370		

Table 5.2. Articles for journals as grouped in Table 5.1.

Contrast that to Table 5.2, which shows article totals year by year and percentages and cumulative percentages for 2016, the most recent full year. Nearly seven of ten 2016 articles are in the 10% most prolific journals, and the bottom 40% of journals account for only 1.2% of articles.

Halves: Journals Grouped in Logical Groupings

Articles	Journals	2017	2016	2015	2014	%	Cum%
1000+	62	61	62	62	58	0.8%	0.8%
500-999	86	86	86	82	76	1.1%	2.0%
250-499	159	156	159	155	134	2.1%	4.1%
125-249	314	302	313	300	270	4.2%	8.2%
63-124	627	600	615	586	527	8.3%	16.6%
32-62	1,016	918	970	911	780	13.5%	30.1%
16-31	1,506	1,217	1,388	1,224	958	20.0%	50.1%
8-15	1,595	1,013	1,301	1,141	867	21.2%	71.2%
4-7	1,089	450	768	701	559	14.5%	85.7%
1-3	1,076	160	424	491	519	14.3%	
Total	7,530	4,963	6,086	5,653	4,748		

Table 5.3. Journals by peak volume, logical groupings

Table 5.3 groups journals logically—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 six rows, with cumulative percentage roughly doubling in each row.

Articles	2017	2016	2015	2014	%16	Cum%
1000+	95,260	80,003	80,190	55,664	26.8%	26.8%
500-999	41,664	42,546	35,148	32,733	14.3%	41.1%
250-499	36,738	39,081	35,483	31,134	13.1%	54.2%
125-249	34,312	37,481	36,910	32,995	12.6%	66.8%
63-124	32,890	35,731	36,078	30,865	12.0%	78.7%
32-62	26,760	29,696	28,260	24,639	10.0%	88.7%
16-31	18,350	20,474	18,237	16,398	6.9%	95.6%
8-15	7,998	9,637	8,919	7,543	3.2%	98.8%
4-7	1,830	2,853	2,797	2,522	1.0%	99.8%
1-3	320	713	823	877	0.2%	
Total	296,122	298,215	282,845	235,370		

Table 5.4. Articles for journals grouped by logical groupings

The 62 largest journals account for 26.8% of 2018 articles; the largest 4% for more than half.

What may be more surprising is that the largest journals appear to be growing in 2017, and growing fairly rapidly: if projections hold up, the 62 largest will be 32% of all 2017 articles, and trends suggest that projections for those journals are probably too low (many seem to be growing month-to-month). Ignore the 62 largest journals and projected 2017 totals would be about 8% lower than for 2016 rather than the less-than-1% drop showed here.

Journals and Articles by *GOAJ2* Levels

Articles	Journals	2017	2016	2015	2014	%	Cum%
600+	124	123	124	121	115	2.0%	2.0%
150-599	393	386	392	381	337	6.4%	8.5%
60-149	800	761	784	743	661	12.9%	21.4%
20-59	1,970	1,706	1,861	1,695	1,402	30.6%	51.9%
1-19	4,243	1,987	2,925	2,713	2,233	48.1%	
Total	7,530	4,963	6,086	5,653	4,748		

Table 5.5. Journal counts by *GOAJ/GOAJ2* levels

This third pair of tables allows for some comparisons with DOAJ-listed journals, as it uses the same row definitions—and since percentages are for 2016, you can make *direct* comparisons with *GOAJ2*.

Articles	2017	2016	2015	2014	%16	Cum%
600+	128,574	113,043	108,399	82,409	37.9%	37.9%
150-599	70,546	76,206	70,789	62,062	25.6%	63.5%
60-149	44,316	48,371	47,238	40,986	16.2%	79.7%
20-59	38,266	42,556	39,456	35,139	14.3%	94.0%
1-19	14,420	18,039	16,963	14,774	6.0%	
Total	296,122	298,215	282,845	235,370		

Table 5.6. Article counts by *GOAJ/GOAJ2* levels

6. Volatility: Countries

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. Additionally, for journals not accessible in 2016 that became accessible in 2017, I did not attempt to assign country or subject. This chapter represents 4,655 journals, 65% of all journals coded A or B—but it includes 79% to 83% of all articles in those journals, the percentage growing in more recent years.
- 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.)
- 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.

While the partial dataset is flawed in general, it may be interesting in looking at volatility: article growth or shrinkage over the four years. That may be especially interesting because casual observation (not realized until too late in the process to formally encode the observation) is that many of the journals in India's massive new UGC list are growing rapidly. I won't replicate and update full tables in *Gray Open Access 2012-2016*; this chapter serves a different and narrower purpose.

Country	Journals	2017	2016	2015	2014
India	2,101	1,587	1,907	1,800	1,402
United States	1,064	652	825	835	763
Nigeria	406	155	248	311	297
United Kingdom	264	147	198	189	194
Canada	203	124	159	155	154
Pakistan	122	72	107	111	91
Hong Kong	75	27	51	57	54
Malaysia	51	29	41	43	38
United Arab Emirates	39	23	35	35	36
Australia	35	24	33	33	27
Bulgaria	28	24	27	28	26
Turkey	28	19	26	28	24
Korea, Rep. of	22	14	19	14	14
Bangladesh	20	16	19	18	19
Kenya	19	1	7	13	18
Singapore	19	17	18	17	17
Romania	15	13	15	13	9
Egypt	13	6	9	11	11
Germany	12	12	12	12	12
Austria	11	11	11	10	2
Iran	10	8	8	10	9
Switzerland	10	10	10	10	9
Subtotal	4,567	2,991	3,785	3,753	3,226
Total	4,655	3,048	3,862	3,835	3,287
India%	45.1%	52.1%	49.4%	46.9%	42.7%

Table 6.1. Countries with most journals in gray OA subset

Table 6.1 shows countries with at least ten gray OA journals (within the 65% subset), sorted by journal count. The most obvious finding: India (which has 284 journals in *GOAJ2*, two of them in APCLand) has *by far* the most gray OA journals, even omitting FA-coded journals and ones

labeled as coming from other countries. In terms of journals actually publishing articles each year, India now accounts for more than half of the subset, up from 43% in 2014.

Articles	2017x2	2016	2015	2014
India	171,771	165,020	145,598	114,704
United States	20,194	17,767	17,024	15,650
Canada	5,664	6,726	7,873	6,619
United Kingdom	4,666	6,125	5,824	4,218
Pakistan	2,960	2,887	3,286	3,958
Romania	1,522	1,253	1,225	1,200
Nigeria	1,478	1,828	2,366	2,471
Bangladesh	1,196	1,565	2,604	3,138
Bulgaria	1,104	1,271	1,824	2,432
Morocco	1,082	1,504	1,191	1,337
Korea, Rep. of	1,070	1,264	1,014	920
Singapore	926	929	1,191	1,562
Austria	864	1,103	1,152	550
Iran	796	850	629	699
Turkey	738	922	918	869
Czech Republic	648	657	88	59
Australia	578	1,035	1,249	897
Russia	552	500	476	413
Japan	490	881	1,057	656
British Virgin Islands	444	457	521	363
Croatia	436	435	374	159
Subtotal	219,179	214,979	197,484	162,874
Total	222,930	219,959	203,293	168,320
India%	77.1%	75.0%	71.6%	68.1%

Table 6.2. Journals with 400 or more (projected) articles in 2017

Table 6.2 looks at articles—and here the picture is even clearer. India accounted for 68% of the 2014 articles in these journals, and that's now up to 77%. Consider the numbers as well (taking into account that most

"United States," "Canada" and "United Kingdom" journals probably aren't, based on the language of the sites).

Country	2017x2	2015	Change
India	171,771	145,598	26,173
United States	20,194	17,024	3,170
Romania	1,522	1,225	297
Korea	468	227	241
Iran	796	629	167
Germany	302	216	86
Russia	552	476	76
Croatia	436	374	62
British Virgin Islands	444	521	(77)
Morocco	1,082	1,191	(109)
Turkey	738	918	(180)
South Korea	602	787	(185)
Singapore	926	1,191	(265)
Austria	864	1,152	(288)
Hong Kong	394	689	(295)
Pakistan	2,960	3,286	(326)
Malaysia	370	721	(351)
United Arab Emirates	248	686	(438)
Japan	490	1,057	(567)
Switzerland	300	925	(625)
Australia	578	1,249	(671)
Bulgaria	1,104	1,824	(720)
Nigeria	1,478	2,366	(888)
United Kingdom	4,666	5,824	(1,158)
Bangladesh	1,196	2,604	(1,408)
Canada	5,664	7,873	(2,209)

Table 6.3. Changes from 2015 to 2017

Table 6.3 may clarify the nature of the volatility. It consists of all countries with at least 200 articles in the gray OA subset in both 2015 and

2017 (projected) and is sorted by the growth or shrinkage, with shrinkage in parentheses (and red if you're viewing this in color). Note that, other than India and the "United States," no country's gray OA publishing grew by more than 300 articles and more than two-thirds shrank, sometimes dramatically.

Articles	2016	GOAJ16	Gray%
India	165,020	29,886	552%
United States	17,767	30,410	58%
Canada	6,726	4,892	137%
United Kingdom	6,125	25,163	24%
Pakistan	2,887	2,798	103%
Nigeria	1,828	540	339%
Bangladesh	1,565	677	231%
Morocco	1,504	114	1319%
Bulgaria	1,271	1,867	68%
Romania	1,253	8,415	15%
Austria	1,103	1,362	81%
Australia	1,035	2,738	38%
Korea, Rep. of	1,264	1,757	72%
Singapore	929	18	5161%
Turkey	922	11,451	8%
Japan	881	899	98%
Iran	850	1,254	68%
Czech Republic	657	2,132	31%
Switzerland	595	1,995	30%
Hong Kong	592	2,725	22%
Russia	500	9,972	5%

Table 6.4. Gray OA articles compared to OAWorld *GOAJ2* articles

Table 6.4 compares 2016 gray OA article counts within this 65% subset (limited to countries with at least 500 articles) to 2016 article counts in

OAWorld, as reported in *GOAJ2*. Adding in APCLand article counts for 2016 would change things, but not as you might expect.

Articles	**2016**	**GOAJ16**	**Gray%**
India	165,020	29,992	550%
United States	17,767	67,190	26%
United Kingdom	6,125	105,972	6%
Australia	1,035	2,910	36%
Korea, Rep. of	1,264	2,313	55%
Singapore	929	114	815%
Japan	881	1,202	73%
Iran	850	1,488	57%
Switzerland	595	40,297	1%
Hong Kong	592	3,304	18%
Russia	500	9,994	5%

Table 6.5. Gray OA 2016 articles compared to all *GOAJ2* 2016, subset

Table 6.5 includes those countries in Table 6.4 that *have* any APCLand journals. The key differences may be that gray OA, even if we accept the stated countries, is a minor part of gold OA in the US and *very* minor part on the UK.

What percentage of India's many gray OA journals are in UGC? That may be something for others to investigate. Are those the journals that are growing rapidly rather than shrinking? Informal observation suggests that they are, at least for the larger journals. Why aren't more Indian journals in *DOAJ*? You'd have to ask the publishers.

7. Volatility: Subjects

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.

Additionally, I did not attempt to add subject names for journals not visible in 2016, so this is once again a subset. It's a larger subset: 79% of all A & B journals, including 93% to 95% of the articles (the percentage *shrinking* slightly in recent years).

I think it might be interesting to see which subjects appear to be growing and which appear to be shrinking.

Since each table fills a page, comments appear after Table 7.2.

Subject	2017x2	2016	2015	2014
Agriculture	7,472	8,321	9,320	9,244
Anthropology	1,056	1,217	1,150	847
Arts & Architecture	710	688	742	756
Biology	7,832	6,897	7,182	6,060
Chemistry	6,042	5,833	6,427	7,209
Computer Science	13,732	16,437	20,343	20,549
Earth Sciences	2,270	2,374	2,745	2,321
Ecology	7,052	8,124	6,861	6,604
Economics	13,540	14,737	14,056	12,650
Education	7,444	7,027	5,825	4,856
Engineering	18,649	19,483	24,303	19,723
History	572	680	862	692
Language & Literature	4,138	3,795	3,470	2,251
Law	1,200	1,382	1,209	1,073
Library Science	460	533	624	441
Mathematics	4,490	5,351	4,929	5,329
Media & Communications	186	319	279	269
Medicine	61,456	59,908	50,039	36,868
Miscellany	43,376	43,753	38,919	28,245
Other Sciences	31,851	28,351	26,752	22,170
Philosophy	24	18	38	32
Physics	702	843	1,048	1,046
Political Science	332	480	477	547
Psychology	620	916	979	563
Religion	272	217	266	200
Sociology	2,594	2,555	3,328	3,210
Technology	6,356	6,123	6,149	6,815
Zoology	3,670	3,026	2,440	1,680
Total	248,098	249,388	240,762	202,250

Table 7.1. Gray OA articles by subject for very large subset of A and B codes

Subject	2017x2	2015	2017% of 2015	2017-2015
Medicine	61,456	50,039	122.8%	11,417
Other Sciences	31,851	26,752	119.1%	5,099
Miscellany	43,376	38,919	111.5%	4,457
Education	7,444	5,825	127.8%	1,619
Zoology	3,670	2,440	150.4%	1,230
Language & Literature	4,138	3,470	119.3%	668
Biology	7,832	7,182	109.1%	650
Technology	6,356	6,149	103.4%	207
Ecology	7,052	6,861	102.8%	191
Religion	272	266	102.3%	6
Law	1,200	1,209	99.3%	(9)
Philosophy	24	38	63.2%	(14)
Arts & Architecture	710	742	95.7%	(32)
Media & Comm.	186	279	66.7%	(93)
Anthropology	1,056	1,150	91.8%	(94)
Political Science	332	477	69.6%	(145)
Library Science	460	624	73.7%	(164)
History	572	862	66.4%	(290)
Physics	702	1,048	67.0%	(346)
Psychology	620	979	63.3%	(359)
Chemistry	6,042	6,427	94.0%	(385)
Mathematics	4,490	4,929	91.1%	(439)
Earth Sciences	2,270	2,745	82.7%	(475)
Economics	13,540	14,056	96.3%	(516)
Sociology	2,594	3,328	77.9%	(734)
Agriculture	7,472	9,320	80.2%	(1,848)
Engineering	18,649	24,303	76.7%	(5,654)
Computer Science	13,732	20,343	67.5%	(6,611)

Table 7.2. Volatility of subjects in gray OA

Table 7.1 simply gives article counts for each year, arranged in alphabetic order. As in *Gray OA 2012-2016*, there are three very large subjects or "subjects"—Medicine, Miscellany and Other Sciences—plus three fairly large subjects (Engineering, Economics and Computer Science).

Table 7.2 is, I believe, more interesting, showing changes from 2015 to 2017. It's arranged by numeric change, from most growth to most shrinkage, although the percentage change also appears.

The three very large "subjects"—each of which covers a multitude of disciplines—are also the only ones with more than 2,000 article growth, although the largest *percentage* growth is in Education and Zoology, the only other subjects with at least 1,000 more articles in 2017 (projected) than in 2015.

But note the bottom half of the table. Nearly two-thirds of the subjects are shrinking in gray OA articles, with some (especially Computer Science and Engineering) shrinking fairly rapidly. That's not a general OA problem, as Computer Science within *DOAJ* journals *grew* slightly from 2015 to 2016 and Engineering grew fairly rapidly from 2015 to 2016. In both cases, it appears that OA publishing is moving from mostly gray to mostly-DOAJ—but it could just be that the gray OA articles are moving to the multidisciplinary megajournals.

.

8. Comments and Conclusions

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

I did a bit of that in *Gray OA 2012-2016* and see no reason to repeat that discussion here.

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). The sheer breadth of the UGC list appears to be increasing publication in Indian gray OA journals, which may or may not be a problem. I suspect a lack of awareness is an issue in some countries, and I suspect that *DOAJ*'s regional and national ambassadors will help rectify this situation. What can India, Nigeria and others learn from South America?

Take away India and the "probably not" countries, and there's not much gray left, as discussed in Chapter 6. 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.

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 and codes.

The dataset is available at https://figshare.com/articles/Gray_OA_2_2014-2017/5500987.

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 9, Whole # 208, 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/civ17i9.pdf