Chapter 1

Idealism and Opportunism

A Gold OA Overview

Open-access (OA) literature is available online to be read for free by anyone, anytime, anywhere, without registration or other hindrance—as long as they have Internet access.

That's the core of OA, and it's a growing factor in scholarly articles (and to a lesser extent monographs). But even as OA grows, confusion as to the amount and nature of OA publishing seems to grow, aided by a lack of clear, concrete information and sometimes by deliberate misinformation.

This report grows from an attempt to determine some of the facts behind the confusion by the simple process of looking at each OA journal site and seeing what was going on.

After a quick refresher course on the basics of OA, this chapter defines the portion of OA being considered here and some of the fundamental issues. An overview of mainstream OA journals shows the extent to which it makes sense to group them into three large and two very small areas—and what constitutes the mainstream.

Chapters 2 through 6 look at OA journals in more detail and with different approaches: journals that charge author-side fees or *don't* charge such fees; journals by volume of articles; some of the economics involved; journals by starting date; and journals by subject category. We'll also look at "journals"—things that have journal names but haven't actually published articles—and other oddities.

Chapters 7 and 8 consider ways to deal with OA journals, suggestions for advising would-be writers, and what libraries and librarians might do to improve the field.

Why should librarians care? Public librarians should care because OA can provide access to research that your patrons may find valuable and that you can't afford to subscribe to. Academic librarians should care because your institution can't keep buying all the journals your community could use and because your library can play an active role in improving the situation (possibly reducing your costs in the long run). All librarians should care because OA means more access to more scholarship for more people.

The Basics

Gold OA consists of journals that make all peer-reviewed articles freely available for online reading as soon as they're published, without requiring fees or registration to read those articles.

Green OA consists of peer-reviewed articles deposited in freely available digital repositories. As currently practiced, green OA may include articles in accepted but not copyedited or formatted form—and may include embargoes.

This report is about gold OA. It's fair to say that "open-access journals" should be synonymous with gold OA. Gold OA does *not* imply author-side charges or article processing fees (APCs): most gold OA journals do not charge such fees.

Because the focus here is on OA journals, this report does not engage in the debate over the relative merits of gold and green.

Some other related terms you may encounter:

- **"Hybrid" journals** are subscription journals that are supposed to offer free access to some articles for which extra, usually very high, fees have been paid. Such journals are *not* included in this report.
- **"Platinum OA"** is an attempt to differentiate gold OA journals with no fees. **"Diamond OA"** is an attempt to define a new category, journals that exist only as overlays on subject archives such as ar×iv. I regard both as needlessly confusing

the issue (Gold OA covers them both), and these terms don't appear in the rest of this report.

- **The *Directory of Open Access Journals*** (*DOAJ*) is an online directory that constitutes the best starting point for consideration of individual journals. All journals discussed in chapters 2 through 5 of this report were in *DOAJ* as of May 2014.
- **Open Access Scholarly Publishers Association** (OASPA) is an association including several of the larger OA publishers. Because OASPA represents less than 25 percent of the journals in *DOAJ*, it is not used as a filter for this report. The few hundred OASPA journals that were *not* in *DOAJ* in May 2014 are discussed in chapter 6 and briefly in this chapter.
- **"Predatory journals"** is a term used by Jeffrey Beall for publishers and journals that he has unilaterally determined to be questionable. (He now uses "possible, potential or probable" to weaken "predatory.") I believe that Beall's lists are irrelevant to any open-minded understanding of OA. Chapter 6 discusses the thousands of journals and "journals" published by publishers and journals on Beall's lists that are *not* in *DOAJ*.
- **Gratis OA** consists of articles that are readable for free online, but possibly no more than that.
- **Libre OA** consists of articles that have at least some additional forms of free usability, ideally including the ability to download, redistribute, use for derivative works, and search or data-mine.

Directory of Open Access Journals (DOAJ)
http://doaj.org

All libre OA is gratis. Although full libre OA is certainly desirable, its presence is such a patchwork that this report does not attempt to distinguish journals that require it and those that allow it.

Journals that impose libre OA will typically state that all contents use a Creative Commons BY license, which requires only attribution for any reuse and is part of the most formal definitions of OA. I have not tracked such statements. When checked in January 2015, *DOAJ* shows less than one-quarter of OA journals using CC-BY as the standard license.

For a deeper discussion of OA, see my 2011 ALA Editions Special Report, *Open Access: What You Need to Know Now*, ISBN 978-0-8389-1106-8.

The Fundamental Issues

This report is not about whether gold OA is worthwhile or whether OA itself is worthwhile. I take those as givens, especially since I'm one of the vast majority of people worldwide who simply does not have access to most scholarly literature *unless* it's OA. This report is also not about whether OA (in all its forms) is growing. That growth is well documented and unmistakable.

What this report does is show the OA journal landscape (or gold OA landscape) as it was in mid-2014 and from 2011 through 2013 and offer breakdowns to make that landscape more understandable.

The OA journal landscape is too complex to show fairly with a few simple facts. For example, "Most OA journals don't impose APCs" is a true statement. But so is "Most OA articles are in journals with APCs." Both statements oversimplify the landscape. Similarly, while an imputed average charge per article is interesting and calculable ($630 in 2013), it's also essentially meaningless.

Some of the questions this report will help answer:

- Is gold OA a significant portion of scholarly publishing—and, if so, how big is it and how fast is it growing?
- How do subject areas differ in terms of gold OA publishing?
- How much money might be involved in gold OA APCs? (That's really two questions: How much do journals charge per article, and how much revenue might journals be gaining from those charges?)
- How many articles are published in a typical OA journal (or, realistically, in various sorts of OA journals)?
- How do OA journals and their policies differ by starting date?
- Beyond major subject areas, do OA journals differ significantly by narrower subject categories?
- How can authors and readers spot questionable journals?

It's possible to give rough single-sentence answers to some of those questions. It's also misleading, with the possible exception of the first question. The simple answers to that multipart question: yes; around 20 percent of refereed scholarly articles in 2013 (and roughly one-quarter of the journals); and around six times as many articles in 2013 as in 2006. But that 20 percent figure involves at least one untestable assumption (that there were two million scholarly articles overall in 2013). By the end of this report you should have a better sense of answers to all of the questions.

Idealism and Opportunism

In the early days of open-access publishing, going back to 1987 and the founding of *New Horizons in Adult Education,* and proceeding at least until 2001,

it's fair to say OA was all about idealism. Groups of people and societies started new online-only journals because they saw gaps in the literature, needs to be met. Few (if any) of the early journals had fees.

During the early years of the new millennium, idealism still dominated the OA landscape, and it still clearly plays a major role: *DOAJ* includes more than 1,400 journals founded since 2009 that do *not* charge fees and many more that charge nominal fees.

But as OA has become more widely known and funding agencies have agreed to support it, opportunism has come into play. In addition to fees required to keep a journal going as it transitions from subscription to open access, there are two new trends: new journals with very high APCs begun by major subscription-journal publishers and new journals begun by small and previously nonexistent publishers because there's money to be made in gold OA. There's also growth in "hybrid" journals, but to date there's little evidence that such journals do much to improve access or stabilize long-term costs.

While chapter 4 includes a more detailed analysis of starting dates in OA journals, a simple set of comparisons (all based on *DOAJ* journals accessible to English-reading people that are actually publishing articles) shows the extent to which opportunism has joined idealism as a basis for OA journals:

- From 1990 through 1999, 507 OA journals began that do not charge fees, while 77 began that do: a free-to-fee ratio of 6.6 to 1.
- From 2000 through 2004, there were 824 new no-fee journals and 144 new fee-charging journals: a ratio of 5.7 to 1.
- From 2005 through 2009, there were 1,322 new no-fee journals and 613 new fee-charging journals: a ratio of 2.2 to 1.
- From 2010 through 2013, there were 1,407 new no-fee journals and 1,181 new fee-charging journals: a ratio of 1.2 to 1.

I think of the period from 2006 through 2012 as the gold rush, and the rush may be declining. It's discussed further in chapter 4.

Opportunism is a loaded term. Publishing *does* cost money, and it's reasonable to believe that large-scale journals are difficult to support entirely on the basis of institutional or association subsidies, without author-side fees or substantial grant funding. And, of course, there's a different balance of idealism and opportunism in a journal charging $100 per article, one charging $1,000, and one charging $5,000 or more. This study doesn't consider how well each journal carries out all the tasks associated with peer review and publishing and whether a journal is sustainable, but those are legitimate questions for journals with no fees or very low fees. That's an exceedingly complex subject since it also raises the questions of whether all the tasks *should* be carried out and at what level.

What fees are reasonable or unreasonable? There's no simple answer to that question. The answers vary based on available government, association, and institutional subsidies (and in-kind subsidies), the size of the journal, the subject area of the journal (some subjects may require much more rigorous peer review than others), and many other factors. I've chosen a breakpoint of $1,000 as one level at which it's reasonable to ask questions about whether a journal should need that much money—but as a single breakpoint, it's arbitrary and certainly wrong in some cases. Based on various initiatives, it's possible to suggest $90 or $500 or $625 or $1,350 as a "justifiable costs per article" point—and those may all be right and wrong.

There are also interesting combinations of idealism and opportunism, including a benign form of opportunism, seeing a gap and filling it. One such gap is the extent to which less-developed nations aren't well served by existing journals; a blend of governmental, cooperative, and private initiatives has yielded hundreds of new OA journals (some with APCs, some without) to provide such service.

Look at that last bullet again: there's *still* a huge amount of idealism in OA publishing. There are more new OA journals *without* fees from 2010 through 2013 than there were in 2005 through 2009—and more than in the 15 years before 2005. The balance may have shifted, but it's not all opportunism by any means. Two clarifications: 218 currently OA journals began before 1990, and there are 196 *very* opportunistic journals in *DOAJ*, ones that almost certainly have APCs but don't state them. Of that 196, 103 started in 2010–2013, and another 53 started in 2005–2009.

A caveat regarding OA journals with no fees: for some of them, "for now" needs to be added as a qualifier, as publishers waive fees for some period in order to increase submissions. That appears to be the situation for 330 journals, or less than 8 percent of the free OA journals in *DOAJ*—and those 330 journals published a total of just under 7,000 articles in 2013, or 5 percent of the total published in no-fee journals. Chapter 5 takes up fees in more detail.

DOAJ and Exclusions

Most of this report deals with a large subset of journals listed in the *Directory of Open Access Journals* as of May 7, 2014. *DOAJ* is the best available resource for OA journals. It doesn't include everything, but it includes almost every OA journal worth considering for readers or authors. Exceptions include brand-new journals and, in the future, journals that publish very few articles each year (fewer than five in any year). Chapter 7 discusses those exceptions.

On May 7, 2014, *DOAJ* included 9,709 journals. I tried to reach more than 8,000 of the journal websites, omitting journals that did not list English as one of their languages, since I would not be able to evaluate those journals. In the end, there were some 700 of those 8,000 that were reachable but didn't have enough English in the interface for me to be able to determine key measures: whether there was an APC (and, if so, how much it was); whether the journal consisted of refereed scholarly articles; and how many such articles were published each year.

The dataset used for this report is 7,301 journals, omitting 2,408 journals. Most journals that publish large numbers of articles are in English or have English as an option. Based on journals that report article-level metadata to *DOAJ*, it appears that roughly 18 percent of articles appear in journals without English as a language option. As a guesstimate, then, article counts in this report may be 18 to 20 percent low. Based on non-English journals that I *could* evaluate and on *DOAJ* figures, it's likely that the vast majority of the 2,407 other journals—70 to 80 percent—do *not* have APCs, so calculations of potential revenue are probably less than 20 percent low.

Article counts are approximate, not only because of missing journals but also because I used approximations in some cases (mostly prolific journals without counting mechanisms) and because what constitutes a refereed article isn't always clear. For smaller journals, where I was manually counting articles, I omitted short communications, book reviews (in most journals), and the like from the counts. For larger journals, such exceptions are rare, and I used shortcuts when available. I'm confident that the picture painted here is accurate in terms of trends and overall patterns; I'm also confident that some of the numbers are *not* accurate down to the last digits. A rule of thumb is probably two or three significant digits—the rest may be approximate. (The exception: potential maximum revenue figures may be *much* higher than reality.)

Other Exclusions

In addition to journals that are opaque to monolingual English readers like me, there are 811 other journals that don't show up in the analysis that follows or in the figures already offered, leaving 6,490 journals that form the basis for this report. The 811 are missing for several reasons:

- **Empty:** 71 journals did not have any published articles between January 1, 2011, and June 30, 2014. Most either explicitly ceased or merged into other journals.
- **Not OA:** 165 journals did not fit the definition of gold OA journals that I chose to use for this study. That includes at least seven consisting entirely of commissioned articles; more than 50 consisting of conference or workshop proceedings; more than a dozen that are magazines (with so few peer-reviewed articles that I couldn't find them); three with embargo periods; at least eight that require subscriptions or otherwise block access; a dozen or so where articles are explicitly *not* peer-reviewed research (*not* including those that do post-publication peer review, which *are* included); a few consisting of monographs or dissertations rather than articles; some journals offering exclusively government agency reports; and at least 40 journals that require registration in order to read articles. Many of these OA publications are worthwhile—but they aren't collections of refereed scholarly articles fully available immediately on publication.
- **Opaque:** I was unable to include 189 journals because the archives were too opaque to count the number of articles in each year without extensive effort. These include more than 50 journals offering only full-issue PDF downloads; roughly 100 journals where the archives did not show dates; and a number where either I couldn't *find* the archive, I couldn't make sense of it, or the archive was so convoluted that I gave up. There was also a couple that insisted on forcing ad windows whenever I took any action on the journal site or in the archive (in addition to at least three journal sites hosting malware, noted in the next section). Originally, there were 295 opaque journals. I checked them directly in *DOAJ* and was able to determine presumed article counts for 106 of them from that source.
- **Unreachable or unworkable:** 386 of the journal sites were unreachable or unworkable. That includes 144 journals where the URL in *DOAJ* yielded a 404 error message; more than 40 that are now parking pages, ad pages, or other things (such as nonroman blogs), including four entirely empty pages; some ten "journals" flagged by Firefox as malicious or that attempted to download malware to my computer; a dozen or so that opened multiple ad windows whenever I took any action; at least 50 with archives or main sites screwed up so badly that they were unusable; and more than 70 that were entirely unreachable (but not 404s) on at least two attempts over different days. That's about a 5 percent failure rate.

A note to readers of *Cites & Insights*: These exclusions may show different and, in at least one case, smaller numbers than the groupings in previous reports (where these would be groups *E*, *N*, *O*, and *X* respectively). That's because I was able to move some journals on further research and modified some criteria.

The bottom line is 6,490 journals that are accessible (at least to some extent) to English-language readers and were reachable on the web; that published at least one refereed scholarly article between January 1, 2011, and June 30, 2014; that publish refereed (peer-reviewed) scholarly articles; that make all such articles freely readable without registration or other impediments as soon as they're published; and that were possible to analyze by date of publication.

Table 1.1. Journals and articles by area

Area	Journals	No-Fee %	Articles	No-Fee %
Mega	4	0%	36,673	0%
Biomed	2,038	47%	128,035	34%
STEM	2,157	60%	141,224	33%
HSS	2,204	87%	52,903	70%
Misc	87	69%	7,375	38%
Total	**6,490**	**67%**	**366,210**	**36%**

Sideshows

Two other sets of OA journals show up in chapter 6, but not in most of this report.

I looked at some 8,000 journals based on Beall's 2014 lists. Of those, 6,948 are discussed in chapter 6, but only 3,256 actually published papers (and were countable) from January 1, 2011, through June 30, 2014.

Just over 400 journals from OASPA members weren't in *DOAJ* as of May 7, 2014. Of these, 308 published articles during the study period. (Quite a few were empty and canceled.) Those 308 published fewer than 6,000 articles in 2013.

Most Beall journals charged fees (as you'd expect: it's hard to be "predatory" if you're not charging anything); most of the 308 other OASPA journals did *not* charge fees.

The Big Picture

Here's the biggest picture, but as you'll see it's somewhat misleading. Of 6,490 gold OA journals publishing just over 366,000 articles in 2013, 67 percent of the journals were free for authors—but those journals published 36 percent of the articles. Theoretically, the journals that *did* charge fees could have taken in around $231 million in 2013, for an average of $1,045 for articles in fee-charging journals or $630 for all articles in 2013. In practice, fee-charging journals almost certainly took in less revenue, given waivers and discounts.

But there's not one big mass of OA journals, all more-or-less the same. There are distinct differences between large subject areas, and there's also extreme variety within each area.

As appropriate, this report deals with journals in three areas—with two small groups of journals handled separately. (The five are exclusive: a journal can be in only one of them.) The three major areas each have roughly the same number of OA journals (2,038 to 2,204); the two other groups (Megajournals and Miscellany) are *much* smaller.

- **Megajournals:** 4 journals each publishing more than 1,000 articles in 2013 (and typically more than 1,000 every year) that cover a wide variety of disciplines. Those four journals account for more than 10 percent of all articles published in 2013, and all charge fees of at least $1,000.
- **Biomed journals:** Those in all aspects of human biology and medicine. Biomed journals publish slightly fewer articles than STEM journals—but Biomed is the only area with a minority of free journals (47 percent), while the percentage of articles in free journals is roughly the same as for STEM (34 percent). This area has *by far* the most potential revenue and is the area in which the gold rush has been most evident.
- **STEM journals:** Those in science (other than human biology and medicine), technology, engineering, and mathematics. This area has the largest number of articles and the greatest disparity between free journals (60 percent) and articles in free journals (33 percent).
- **HSS journals:** Those in the humanities and social sciences. The largest number of journals, *by far* the smallest potential revenue of the three areas, and the area in which free publishing dominates both journals (87 percent) and articles (70 percent).
- **Miscellany:** 87 journals that either cross too many disciplines to fit into one of the three areas or that couldn't reasonably be assigned to one of them. Fewer than 7,400 articles, with a tiny amount of potential revenue.

Note that in tables and discussions the terms *free* and *no-fee* are interchangeable, as are *fee* and *APC*—and *free articles* means articles in journals that don't charge fees.

Table 1.1 shows the overall picture for articles published in 2013 and journals discussed in this study. (Of the journals that were free when checked but might impose APCs later, 161 are in Biomed, 139 in STEM, 26 in HSS, and three in Miscellany.)

Table 1.2 shows the *potential* revenue in each area—how much would have been taken in if every 2013 article resulted in the full APC or other charge. **$/Article (APC)** divides that amount by the number of articles in fee-charging journals, while **$/Article (All)** divides that amount by the total number of articles.

Table 1.2. Potential revenue by area

Area	Potential Revenue	$/Article (APC)	$/Article (all)
Mega	$49,637,565	$1,354	$1,354
Biomed	$114,440,937	$1,460	$894
STEM	$59,624,766	$681	$422
HSS	$6,419,931	$439	$121
Misc	$605,987	$176	$82
Total	**$230,729,186**	**$1,045**	**$630**

Note the huge differences here: Biomed articles overall cost more than twice as much as STEM articles, which in turn cost more than three times as much as HSS articles.

That's the big picture. The rest of this report fills it out and adds other measures, exploring the diversity of gold OA journals.

Acknowledgments

Thanks to Stephanie Willen Brown, Cameron Neylon, and John Dupuis for reviewing the draft of this report and offering great advice.