Dramatic Growth of Open Access March 31, 2013

Commentary: Revised April 4, 2013

Correction and clarification: the number of articles added by OASPA in 2012 was 81,780, not 25,788. This would place CC-BY growth in the same ballpark as arXiv, but still less than Highwire Press and much less than repository growth.  An anonymous commenter reports the split of DOAJ articles by license type. I am not able to replicate the search and do not have contact information for this person. If anyone can explain to me how I can do a DOAJ search by license without having a specific term that would be appreciated, as would contact information for the anonymous commenter. The chart above and data below have been updated to reflect this correction.

Clarification: the BASE documents number is used as a very rough surrogate of open access growth through repositories. Not all works in repositories are open access, and not all are scholarly works. However, I argue that the sheer size of the growth strongly suggests strong growth of open access in repositories. If only 1% of the 9.2 million document growth of BASE reflects open access scholarly peer-reviewed journal articles, that would still be close to 100 thousand articles added in 2012. This consideration is described in my [rationale and method](https://docs.google.com/document/d/1MV4_UDyerFF5fXkdkcyf1SPoePX0sBNeAecxyoNvNy0/edit?hl=en_US). Another important point is that open access repository growth isn't just about scholarly articles - this growth reflects a lot of grey literature, for example, reports and theses that are valuable scholarly works that generally received very little dissemination in previous years, increasingly research data, and the historical works are valuable primary documents for many scholars. This is a whole range of growth of open access materials that is taking place largely within the repository movement, and very little in open access publishing.

Thanks very much to [*Jyrki Ilva*](http://blogs.helsinki.fi/digikirjasto/) and "Alf" the anonymous commenter for this most welcome peer review. Can this now be considered a peer-reviewed blogpost?

**Highlights**

This issue features a comparison of open access growth including [CC-BY article growth figures supplied by OASPA](http://oaspa.org/growth-in-use-of-the-cc-by-license-2/). In brief: for every CC-BY article addition tracked by OASPA, repositories around the world add 359 documents as found by a BASE search, DOAJ adds 10 articles that are not CC-BY licensed (90% of DOAJ article growth), arXiv and SSRN each add 3 documents, and the Internet Archive adds lots of texts, movies, sound recordings and concerts. Recent research suggests that CC-BY is the preference of a small minority of scholars.

The top 10 growth figures by percentage for both this quarter and the past year are presented. Looking at percentage growth brings out substantial growth in initiatives with smaller numbers. Note that smaller numbers are not necessarily less significant. One open access funding agency mandate can mean free access to tens or even hundreds of thousands of articles, for example. Open access mandates are high on the list of percentage growth figures, including 26 funding agency OA mandates this quarter for a total of 80 and a growth rate of 48%. The Directory of Open Access Books is growing up leaps and bounds, or to be more specific added 13 publishers and 135 books this quarter. The usual suspects (Directory of Open Access Journals, PubMedCentral, and BASE) continue to rank highly on percentage comparisons. Highwire Press added a total of 20 totally free sites this past year for a total of 71, an impressive sixth place (not bad for an initiative that isn't focused on open access).

Kudos to [DOAJ for hitting the 1 million article milestone](http://www.doaj.org/doaj?func=news&nId=44&uiLanguage=en). Bjork, Laakso, Welling and Paetau have issued a preprint of another major open access growth study, the [Anatomy of Green Open Access](http://www.openaccesspublishing.org/apc8/index.html), finding that the coverage of all journals articles as green open access is currently at 12%. Suber has posted [additional figures and analysis](https://plus.google.com/u/0/109377556796183035206/posts/iSR2spVGFUL) and updated the [open access by the numbers](http://oad.simmons.edu/oadwiki/OA_by_the_numbers#OA_journals_.28Gold_OA.29) section of the Open Access Directory. New this issue is the amazing 281 billion web pages of the Internet Archive.

Full data, a word version of this commentary and jpg of the chart above are available in [SFU SUMMIT](http://summit.sfu.ca/item/12645).

**Comparisons and CC-BY**

This issue features a comparison of some of the top open access growth figures, by number and percentage for this quarter and the past year.  This comparison is inspired by some of my colleagues in the open access movement who equate open access with the Creative Commons - Attribution Only license and appear to believe that this is becoming a default for open access, as illustrated by the Open Access Scholarly Publishers Association's odd release of CC-BY article data (this is odd because not all members use this license, and OASPA has not released data for other licenses used by members).

The chart above shows number of works added from March 31, 2012 to March 31, 2013 by source, limited to the top 9 sources of those tracked in this series plus the OASPA CC-BY figures. On the left we see that the growth of documents in repositories encompassed by a Bielefeld Academic Search Engine search dwarfs all other figures, with over 9.2 million documents added this past year. I include figures from the Internet Archive even though most are not scholarly works because this resource and the phenomenon of digitization and sharing of current and past works of various kinds is worth tracking, with this initiative being just one example. This past year IA added over 1 million texts, half a million movies, 360,000 audio recordings and close to 14 thousand concerts. The number of articles searchable at article level in the Directory of Open Access Journals grew by over 270,000. The Highwire Press free articles collection grew by over 130,000. arXiv grew by over 86,000 documents, and the Social Sciences Research Network by over 65,000 full text papers. OASPA tracked growth of over 25,000 documents.

Another way to compare these figures for CC-BY aficionados: for every CC-BY paper added by OASPA, the following were added by these services (corrected April 4, 2013):

* 113 BASE documents
* 13 Internet Archive texts
* 7 Internet Archive movies
* 5 Internet Archive recordings
* 3 DOAJ articles searchable at the article level
* 1.6 Highwire Press free articles
* 1 arXiv documents

Other perspectives

Of the 271,715 articles searchable at the article level added by DOAJ, the 81,780 CC-BY articles as per OASPA data represents 30%. Another way to express this is that 70% of the articles added to the DOAJ article search in 2012 are NOT CC-BY. (Updated April 4, 2013).

Most of the journals participating in Highwire Press are fairly traditional scholarly society journals. These journals are providing access to free articles at a rate **1.6** times higher than the CC-BY group. (Corrected April 4, 2013 - was 5).

Note also that research indicates that [CC-BY is the preferred license of only a small minority of scholars](http://poeticeconomics.blogspot.ca/2013/04/attitudes-and-values-regarding.html) (5-10%).

Highest growth this quarter by percentage (top 10 plus 2 due to 3-way tie for 10th place)

1. Funding agency open access mandates (ROARMAP): 26 mandates added this quarter for a total of 80 (48% growth this quarter).
2. Publishers participating in the Directory of Open Access Books: 13 added this quarter for a total of 48 (37% growth this quarter).
3. Proposed open access mandates (ROARMAP): 4 added this quarter for a total of 27 (17% growth this quarter).
4. The number of journals in PubMedCentral providing immediate free access: 177 more this quarter for a total of 1,203 journals (17% growth this quarter).
5. The Internet Archive added 616,132 texts this quarter, bringing the total to over 3.3 million (16% growth this quarter).
6. Highwire Press added 8 completely free sites this quarter, bringing the total to 71 (a 13% increase this quarter).
7. The Directory of Open Access Books added 135 books this quarter, for a total of 1,394 (11% increase this quarter).
8. The Directory of Open Access Journals added 100,097 articles searchable at article level this quarter, for a total of 1,055,817 (10% increase this quarter).
9. The number of journals in PubMedCentral depositing selected articles: grew by 166 journals to a total of 2,064 (9% growth this quarter).
10. The number of documents included in a Bielefeld Academic Search Engine search grew by over 3 million to a total of more than 43.5 million (7% growth this quarter).
11. The number of journals searchable at the article level in the Directory of Open Access Journals grew by 292 for a total of 4,539 (7% growth this quarter).
12. The Internet Archive added more than 75 thousand movies this quarter for a total approaching 1.2 million (7% growth this quarter).

Social Sciences Research Network - paper downloads in the last 30 days - up 33% (likely a fluke as Dec. 31 would have reflected the slower December period).

Highest growth in past year (March 31 - March 31) by percentage. The OASPA CC-BY article figure is added for illustration purposes although this data is not part of the dramatic growth series.

1. Multi-institutional funding agency open access mandates: 3 added in the past year for a total of 4. (300% growth).
2. Internet archive moving images (movies): 556,115 added in the past year for a total of close to 1.2 million (88% growth).
3. Funding agency open access mandates: 28 added in the past year for a total of 80 (54% growth).
4. **CC-BY articles: 81,780 added in the past year for a total of over 250,000 (46% growth). (Based on data supplied by OASPA). Corrected April 4, 2013.**
5. Proposed open access mandates: grew by 8 in the past year for a total of 27 (42% growth).
6. Highwire Press completely free sites: grew by 20 in the past year for a total of 71 (39% growth).
7. # of journals participating in PubMedCentral with immediate free access: 324 added this year for a total of 1,203 (37% growth).
8. Directory of Open Access Journals # of articles searchable at article level: grew by 271,715 this year for a total of over 1 million (35% growth).
9. # journals in PubMedCentral with all articles open access: 230 added this year for a total of 976 (31% growth).
10. Social Sciences Research Network - paper downloads in the last 12 months: grew by 2.6 million for a total of 11.4 million (30% growth).

**The numbers**

 [Directory of Open Access Journals](http://www.doaj.org/)

* 8,847 journals
* 328 journals added this quarter - growth rate of over 3 per day
* 4,539 journals searchable at article level - 292 added this quarter
* over 1 million articles searchable at article level - over 100,000 added this quarter - growth rate of over 1 thousand articles per day

[Directory of Open Access Books](http://www.doabooks.org/)

* 1,394 academic peer-reviewed books; 135 added this quarter (more than 1 per day)
* 48 publishers; 13 added this quarter

[Electronic Journals Library](http://rzblx1.uni-regensburg.de/ezeit/index.phtml?bibid=AAAAA&colors=7&lang=en)

* 39,227 journals that can be read free of charge; 1,422 added this quarter (15 per day)

[Highwire Free](http://highwire.stanford.edu/lists/freeart.dtl)

* 2.3 million free articles; over 100,000 added this quarter - over 1 thousand per day
* completely free sites: 71: 8 added this quarter
* sites with free back issues: 284: 2 less this quarter

[OpenDOAR](http://www.opendoar.org/)

* 2,269 repositories: 16 added this quarter

[Registry of Open Access Repositories](http://roar.eprints.org/#repositories)

* 3,379 repositories: 39 added this quarter

[Bielefeld Academic Search Engine](http://www.base-search.net/)

* 43.5 million documents
* 3 million documents added this quarter - over 30 thousand per day

[PubMedCentral](http://www.pubmedcentral.nih.gov/)

* 2.6 million items (not updated regularly)
* 1,487 journals actively participating (23 added this quarter)
* 1,203 journals with immediate free access (177 added this quarter or 2 per day)
* 976 journals with all articles open access (83 added this quarter or 1 per day)

[arXiv](http://arxiv.org/)

* 832,859 documents
* 23,010 documents added this quarter (over 250 per day)

[Social Sciences Research Network](http://papers.ssrn.com/sol3/DisplayAbstractSearch.cfm)

* 385,838 fulltext papers
* 13,066 added this quarter (145 per day)

[Open Access Mandate Policies (ROARMAP)](http://www.eprints.org/openaccess/policysignup/)

* 36 sub-institutional (2 added this quarter)
* 80 funder (26 added this quarter)
* 165 institutional (2 added this quarter)
* 4 multi-institutional
* 100 thesis
* 385 total (32 added this quarter)
* 27 proposed

[Internet Archive](http://archive.org/index.php)

* 281 billion web pages (new)
* 1.1 million movies: 75 thousand added this quarter (over 800 per day)
* 114 thousand concerts: 3,789 added this quarter
* 1.5 million audio recordings; close to 100 thousand added this quarter
* 4.3 million texts; 616,132 added this quarter (close to 7 thousand per day)

 This post is part of the [Dramatic Growth of Open Access series](http://poeticeconomics.blogspot.ca/2006/08/dramatic-growth-of-open-access-series.html).



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License](http://creativecommons.org/licenses/by-nc-sa/3.0/deed.en_US).