

Discovering Scholarship on the Open Web:

Communities and Methods

Joan Fragaszy Troyano, Project Director
Jeri Wieringa, Research Assistant
Project supported by the Alfred P. Sloan Foundation

Online publications that aggregate content from a wide variety of sources have become increasingly valuable to readers and publishers. The academy, however, is still unsure how to efficiently identify, collect, survey, evaluate, and redistribute the valuable scholarly writing published both formally and informally on the open web. Fortunately, some scholarly communities are developing methods to draw attention to upcoming work in their fields.

This report outlines the current state of the aggregation, curation, evaluation, and distribution of scholarship on the open web. We describe the primary types of websites where open collections of scholarly work can be found, specifically repositories, aggregators, curated content, and forums for post-publication review. We suggest an eight-point rubric for analyzing similar sources of web-published scholarship. Finally, we offer an annotated bibliography of outlets for scholarly communication on the open web.

Table of Contents

- I. Introduction
- II. Aggregating and Curating Scholarly Content on the Web
 - o Repositories
 - o Collections of Aggregated Content
 - o Curated Content
 - o Post-publication Review and Community Discussion
- III. Conclusion: Moving toward a Rubric
 - o Rubric for Scholarly Communication Outlets on the Open Web
- IV. Appendices
 - o Appendix A: Sites Reviewed
 - o Appendix B: Tools And Services Examined
 - o Appendix C: Scholarly Communication Projects on the Open Web

I. Introduction

In the third decade of the web, scholars have become used to — even if they have not entirely embraced — accessing scholarly writing online. Formally published original research appears in open access journals, or through institutional subscriptions to individual journals or databases such as [JSTOR](#) or [EBSCO](#). At the same time, research centers and individual scholars are sharing the equally important and much more timely “gray literature” of grant reports, research notes, writing of various lengths, presentations, and pre-print and open access versions of their publications through their own web presence, on privately-held social networking sites like [academia.edu](#), [figshare.com](#), or [researchblogging.org](#). Increasingly, they offer, or are required, to place published works in their own institutional repository. The burden of discovering new work, both formally published and gray literature shared in these ways, however, falls to the researcher who must monitor the field by assembling and aggregating their own web sources of information from repositories, blogs, and [Twitter](#).

Academics, along with other research and writing professionals working with large amounts of information, want to tame the firehose of information. In fact, everyone from individual readers to media giants and academic communities seek methods to identify, evaluate, collect, and highlight the information desired without getting overwhelmed or sidetracked.

Online publications that aggregate a large number of sources of information have become increasingly valuable to readers and publishers. Sites with an additional layer of editorial oversight — from human editors or algorithms — that curate the aggregated content are even more valuable to readers than simple aggregation. This white paper reviews the state of aggregated and curated scholarly content on the open web.

To some degree every reader is his or her own aggregator and curator by selecting the sources they read regularly and the links that they follow. The more committed create personalized aggregation feeds through various subscription methods, from email digests to Twitter lists and RSS. A “reader” or bookmark program like [Instapaper](#) or [Pocket](#) (formerly [Read It Later](#)) can make it easier to see everything within the same interface and quickly determine what is new. Web services to help users stay abreast of their own social network have appear regularly in the past few years, such as [Flipboard](#), [Summify](#), [paper.li](#), and [Tweeted Times](#), that publish newsletters and magazines based on user criteria. Joining the free filter services like [news.me](#) are subscriptions like [Percolate](#) or proprietary software such as [Fever](#). But many individuals still rely on the old-fashioned method of going directly to web pages on a regular basis — hourly to weekly — for information, work, or pleasure.

Media entities such as the [Daily Beast](#) and [Huffington Post](#) made their name by reliably providing their readers with content that had the most buzz. [Techmeme](#), [Technorati](#), and [MetaFilter](#) have been successful by aggregating information, creating methodologies to curate the content, and attracting, even amassing, attention from a large number of readers. The aggregation and curation of online content is an extension of the traditional activities of the media, and so they have led the way in developing technological methods and workflow to do so.

Scholars have the skills required to curate content. Central to graduate training in every discipline is the ability to identify and survey a vast amount of information, evaluate its usefulness, and highlight the most important components. Unfortunately, academic organizations and non-profits have not been able to replicate the reach or success of media entities because the technical means are proprietary, and/or the necessary editorial staff is difficult to sustain. The importance of communicating research and advancements in the field to academics, policy makers, and the public is great, and necessitates an improved system.

As a whole the academy is still unsure how to navigate, identify, survey, evaluate, and highlight the large amount of scholarly writing, research, and publications available on the open web. Fortunately, some scholarly communities are developing methods to direct attention to new and valuable work and conversations in their field. This report outlines the current state of the aggregation, curation, evaluation, and distribution of scholarship on the open web.

We describe the primary types of websites where open collections of scholarly work can be found, specifically repositories, aggregators, curated content, and forums for post-publication review. We suggest an eight-point rubric for analyzing similar sources of web-published scholarship. Finally, we offer an annotated bibliography of outlets for scholarly communication on the open web.

II. Aggregating and Curating Scholarly Content on the Web

Collections of scholarship on the open web fall into one or more of the following genres: the traditional journal reproduced online, repositories of articles or data, collections of aggregated content for the reader to peruse; curated content influenced by editors, and post-publication review sites meant to evaluate research.

Scholarly content on the open web is most likely to be found in a repository or post-publication review site. The sciences and social sciences in particular have embraced these formats, with the humanities lagging behind in every category. In all fields, the greatest need is for the creation of reproducible, efficient, and effective aggregation and curation methods.

Repositories

Repositories offer scholars a central location to find and deposit their work. The best examples of repositories of scholarly content are the very active and established [arXiv.org](http://arxiv.org) and [Social Science Research Network \(SSRN\)](http://www.ssrn.com) which attract pre-publication papers, working papers, and gray literature such as conference papers. For readers seeking new work in the fields of physics, mathematics, computer science, quantitative biology, quantitative finance, statistics, and an array of the social sciences, [arXiv.org](http://arxiv.org) and [SSRN](http://www.ssrn.com) are the first place to look, rather than printed journals.

Not surprisingly, given that [arXiv.org](http://arxiv.org) and [SSRN](http://www.ssrn.com) have replaced printed journals as the source for new work, both sites were initiated and developed in the 1990s outside of professional organizations (in fact, [SSRN](http://www.ssrn.com) is a privately-held corporation). Currently most scholarly organizations publish newsletters and journals that require a paid subscription rather than developing their own open access repositories or curating topical content found elsewhere on the web. Scholarly organizations may begin to follow the lead of the American Folklore Society, however, which has initiated [Open Folklore](http://openfolklore.org) to make new work and making previously published content available and accessible at no charge.

In the quantitative sciences and social sciences, then, sharing pre-publication material in open repositories is common, and there are established locations for uncovering new work. In the life and biomedical sciences, however, formal publication remains the priority. The open access journals of the [Public Library of Science \(PLOS\)](http://pubs.plos.org) have become an important outlet for these fields for this reason, because these journals are not limited in the number of articles they publish and so are able to bring more research to light at a faster rate than print journals. (They do, however, charge article-publishing fees in order to release as open access.) [PLOS](http://pubs.plos.org) also has initiated the [PLOS One](http://journals.plos.org/plosone/) journal as a hybrid repository-journal because it provides a basic and fast review of the integrity of the content to ensure quick publication. [PLOS One](http://journals.plos.org/plosone/) requires the readers to determine the significance of a piece, and includes an online forum to encourage discussion of the content. The humanities and liberal arts lack an equivalent repository for pre-print materials, gray literature, or broad disciplinary journals for quick and open publication.

Successful repositories rely on users to validate their existence through the provision of content and interest in accessing the content. When acknowledged and accepted by disciplinary communities, as they are in the sciences and social sciences, they are extremely useful. They do not, however, provide context for understanding, and only minimal evaluation of the content available from metrics such as views and download history, or user ratings. Ultimately, the user of a repository must rely on their own searching and evaluation skills to determine the value of the repository and the content.

Collections of Aggregated Content

There are fewer examples of informal scholarly communities who have organized group efforts to identify and aggregate new work. The best of these is [Science Seeker](#), which aggregates and categorizes work from hundreds of blogs. Readers can visit the main page to see all recent posts, sort by category, or subscribe to topics of interest in order to stay on top of the most recent writing on science blogs. For the digital humanities community, [2cultures.net](#) syndicates blog posts from 100 different sources, with no obvious editorial hand or commentary from the single organizer Craig Bellamy; it is simple republishing by linking.

Curated Content

Many successful publications on the general web aggregate and curate information through a heavy editorial hand, with [Daily Beast](#) and [Huffington Post](#) leading general news category. [The Browser](#) has joined [Arts & Letters Daily](#) as a source of high-quality long-form writing through reliance on multiple editors. Recently the [Walker Art Center](#) in Minneapolis began to link to a curated selection of art news around the world from their home page. The Walker considers their web presence to be a publication and staffs it accordingly, and may prove to be the model for museum websites in the coming years. The few examples of sites that curate scholarly material from the open web are discussed below.

Two general web publications, [Techmeme](#) and [Technorati](#) and all their affiliated sites, have had success using algorithmic methods of filtering as a form of curation. [Techmeme](#) uses algorithms to gather and prioritize news stories from around the web, with categorization and publishing overseen by human editors. [Technorati](#) provides a ranking of “authority” for blogs in topical categories through algorithms based on crowdsourcing: the number of links to the site, and the number of times it is referenced by other sites. Like [Techmeme](#), [Technorati](#) is produced by a private company, employs editors, and is supported through ad revenue. The reliance on paid development and editorial staff by these two successful sites suggests the two biggest challenges to recreating these methods in the non-profit, academic environment: lack of time and money.

Some individual scholars have taken it upon themselves to curate and distribute information relevant to their field. Jay Rosen of New York University, for example, broadcasts news and commentary on developments in journalism and the media industry to more than 110,000 followers on Twitter and readers of his blog, [PressThink](#), and [Tumblr page](#). Rosen’s audience is not surprising for a field like media studies, with its contemporary focus and interaction with current events. It is a major effort, however. Even Rosen, who personally monitors more than 1,000 blogs and Twitter feeds, including a number of aggregation and curation sites, does not have a simple method to stay on top of all this information; [he intentionally and repeatedly visits his sources of information](#), rather than pulling it into one single platform to monitor, review, link, draft, and publish.

There are three notable and recent experiments in the curation of scholarly material coming from disparate groups and fields: [PLoS Hubs](#), [Cléo’s Radar](#), and [RR-CHNM’s Digital Humanities Now](#). [PLoS Hubs](#) are topically-focused

sites on Biodiversity and Clinical Trials that are prepared by several expert editors. Each Hub reproduces articles from across [PLoS](#) journals along with additional content and linked data to provide more complete context in which to understand a topic, such as biodiversity. The French organization Cléo, host of the [Hypotheses](#) blogging platform and [Open Edition](#) press among other projects, has been experimenting with the [Radar](#) project in which eight editors survey their respective fields and provide links to the most important information. Finally, RR-CHNM has been running [Digital Humanities Now](#) for more than two years, and in its third iteration has moved towards a greater reliance on editorial evaluation of aggregated information, while continuing to experiment with crowdsourcing content and selection.

Although there are different disciplinary values for the projects that curate scholarly content, behind each of these curated projects are human editors with their own methodologies for surveying large amounts of information, identifying connections, assessing quality, and selecting content for publication.

Post-publication Review and Community Discussion

Finally, there are very active sites for post-publication peer review. For the biomedical sciences, Faculty of 1000 ([F1000](#)) is a subscription-based digital-only publication that provides post-publication peer review. Vetted scientists called “Faculty Members” review, rank, and draw attention to most interesting articles in their research areas. Economists can look to the open access [EconJournalWatch](#), a post-publication review site, for original analysis, rather than highlights, of previously published content. For legal studies, the open access [JotWell](#) provides a similar service: reviews of recently published journal articles meant to help readers stay on top of their field and aware of developments in the broader discipline. Once again, the sciences and social sciences lead the way in the number and success of post-publication review sites.

Whether a site functions as a repository, aggregator of content, curated collection, or post-publication review site, there are a variety of ways to acquire content. Most often a repository requires direct submission from a user or active oversight from an administrator. Sites that aggregate content, like [Science Seeker](#) or [2cultures.net](#), have passive submission through feeds and social media. Curated sites, such as Cléo’s Radars and [Digital Humanities Now](#) may rely on passive submission, but have active editors. The post-publication review sites are completely dependent on authors providing the reviews.

The editorial oversight on these sites ranges as well. There may be a monarchical single editor, like Jay Rosen, an oligarchy of multiple editors with specific areas of oversight, like [F1000](#), an active democracy of readers directly winnowing out the content they want to see by ranking articles in PLoS journals, or a passive democracy of readers supplying metrics such as the download and citation numbers visible in [SSRN](#). No matter the type of editorial oversight, the criteria for which scholarly content gets published on curated sites are not that different from criteria for traditional publication: quality of ideas and relevance to a field.

The methods used to determine the quality and relevance of scholarship may differ, however. The most traditional option is editorial oversight by content experts. The web makes two additional types of criteria possible: metrics based on use of key words, length, or citations; and metrics based on reader interaction through views, downloads, user rankings, or number of comments, citations, and social media.

For the moment, the reproduction of quality information is dependent on the critical eye, and time, of dedicated and knowledgeable editors. The challenge is to automate as much as possible the filtering out of non-essential information, to make the editors’ jobs easier, and to increase opportunities for crowd-sourcing the identification and evaluation of quality scholarship.

III. Conclusion: Moving toward a Rubric

There are a handful of notable projects that aggregate and curate scholarship on the open web. The sciences and quantitative fields are the most established in this landscape. arXiv.org is the best example of community-organized and supported repository. [Science Seeker](http://ScienceSeeker.org) shows that there are a great number of scholars that will independently provide insight and results from their research, and that aggregating scholarly work makes communication easier within, and across fields. The audience for post-publication peer review is so great that it even can be lucrative, as [F1000](http://F1000.org) proves.

In each of these cases, it is informal communities of scholars, or privately-held companies, that provide the foundational work of identifying and cultivating the scholarly work on the web. Scholarly societies for the most part have not taken the lead in creating open repositories for new research or gray literature, or to aggregate and/or curate content from their members.

Clearly, scholarly communities will do the work required to gather and share knowledge with each other. What is needed is a thorough evaluation of the extent to which these projects have impacted scholarly communication and met the needs of their respective communities.

In addition, a free, replicable, efficient, and effective method to automate the aggregation and curation of scholarly content is needed. A customizable tool for the algorithmic methods of discovery and evaluation, along with a place for internal or crowd-sourced editorial conversation, will be necessary for any group that wishes to collect, select, and publish scholarly content from the web. It is these needs that will inform the creation of the forthcoming PressForward suite of plugins for the [WordPress](http://WordPress.org) publishing platform.

On the next page we offer a rubric for analyzing new and existing outlets for scholarly communication on the open web.

Based on our extensive examination of sites and services (please see the Appendices), we offer the following rubric for analyzing new and existing outlets for scholarly communication on the open web.

Rubric for Scholarly Communication Outlets on the Open Web

1. Function of Site
 - Repository of content
 - Aggregated content
 - Curated content
 - Community discussion
 - Traditional journal
 2. Type of Content
 - Original content
 - Previously published content
 - Post-publication commentary and discussion
 3. How Content Submitted
 - Solicited and developed by editors
 - Direct submission by user
 - Passive submission through feeds/social media
 4. Who Controls Content
 - Monarchy: Single Editor
 - Oligarchy: Multiple Editors
 - Democracy: Readers
 5. How Published Content Is Determined
 - Editorial oversight
 - Algorithm
 - Active reader participation
 - Passive social media
 6. How Content is Selected for Review
 - Traditional editorial criteria of identifiable quality
 - Amount of interaction (comments, citations, links, alt-metrics, etc.)
 7. How Distributed/Consumed
 - Static website
 - Twitter feed
 - RSS feed
 - Email notifications
 8. Business and Sustainability Model
 - Advertisements
 - Membership or reader-supported
 - Author-supported
 - In-kind labor
 - Paid staff
-

IV. Appendices

Appendix A: Sites Reviewed

1. [MetaFilter](#)
2. [TechMeme](#)
3. [Technorati](#)
4. [Microsoft's Academic Search](#)
5. [PLoS One, Blogs, Hubs](#)
6. [Scientific Reports](#)
7. [arXiv.org](#)
8. [viXra.org](#)
9. [dspace.org](#)
10. [SSRN](#)
11. [Research Papers in Economics](#)
12. [Philpapers](#)
13. Phyg (no longer active)
14. [Open Folklore](#)
15. [Journalists Resource](#)
16. [Merlot](#)
17. [Jotwell](#)
18. [Faculty of 1000](#)
19. [Econ Journal Watch](#)
20. [Third Reviewer](#)
21. [Polymath Blog](#)
22. [Journalists on Twitter](#)
23. [Science Seeker](#)
24. [Science Blogs](#)
25. [ScienceBlogs](#)
26. [Scientopia](#)
27. [PressThink](#)
28. [Marginal Revolution](#)
29. [The Volokh Conspiracy](#)
30. [Cliopatria](#)
31. [BLDGBLOG](#)
32. [MediaCommons](#)
33. [The Art of Theory](#)
34. [academicblogs.org](#)
35. [hypotheses.org](#)
36. [Scholarly Exchange](#)
37. [The Conversation](#)
38. [Nature Precedings](#)
39. [science.io](#)
40. [up2date](#)
41. [2cultures](#)
42. [Immanent Frame](#)
43. [Peer Evaluation](#)
44. [Sympose](#)
45. [Mathoverflow](#)
46. [Arts and Letters Daily](#)
47. [The Browser](#)

48. [Poynter Institute](#)
49. [Walker Art Center](#)
50. [Nieman Labs](#)
51. [Digital Humanities Now](#)
52. [academia.edu](#)
53. [figshare.com](#)
54. [researchblogging.org](#)

Appendix B: Tools and Services Examined

The following tools are used for aggregation, curation, and creation of scholarly content on the web. Their services may complement, or be integrated into, the PressForward suite of plugins for the [WordPress](#) publishing platform.

1. [Anthologize](#)
2. [Scholastica](#)
3. [Digress.it](#)
4. [Annotum](#)
5. [Postrank](#)
6. [Feedrank](#)
7. [Percolate](#)
8. [suu.sh](#)
9. StrawberryJ.am (no longer active)
10. [paper.li](#)
11. [Flipboard](#)
12. [TweetedTimes](#)
13. [Yahoo pipes](#)
14. [Summify](#)
15. [Bottlenose](#)
16. [Scalar](#)
17. [Fever](#)
18. [Storify](#)
19. [ThoughtMesh](#)
20. [news.me](#)
21. [Scoop.it](#)
22. [Zementa](#)

Appendix C: Scholarly Communication Outlets on the Open Web

- Repositories
- Aggregation Projects
- Curation Projects
- Post-Publication Review and Community Discussion
- Options for Traditional Publication on the Open Web

Repositories

1. [arXiv.org](#) is a non-profit, online repository for papers in the sciences, particularly physics, math, computer science, and quantitative fields. The site allows registered users to submit articles of original research that are pre-print, works in progress, completed, or already published. Articles may be revised and resubmitted multiple times by the author. Before submitting articles, registered users must be endorsed by existing members/authors. This is not a peer review process, but a way to ensure that the author will submit appropriate content. The article abstracts and author information are viewable on the site and the entire document is available for download in PDF, PostScript, DVI, or as Source files. The author's submission email is viewable by registered users. arXiv.org is a non-profit organization housed in the Cornell University library system and also partially funded by the library. In order to meet their budget of approximately \$600,000, arXiv.org has requested financial support from the most frequent user institutions.
2. [MERLOT](#) is a digital repository for teaching resources in multiple disciplines. There are a wide variety of materials available, including tutorials, collections, learning objects, assignments, case studies, and textbooks. A portion of the material is peer reviewed by site members through a process overseen by the editorial board. Registered members can interact with the site by contributing comments or new materials, creating a personal collection, or providing peer review upon assignment from the editors. The materials are posted as finished products; revisions are not grouped. There is a star system for ranking, both for peer reviewers and member comments. MERLOT is sustained by the University of California system, other partners in higher education and philanthropy.
3. [Nature Precedings](#) is an open access repository for pre-publication research in the sciences and medicine. Manuscripts, posters, and presentations may be submitted in Word, Powerpoint, or PDF formats and then are approved by editors for availability on the site, although they are not peer-reviewed. Site members provide feedback on the works by commenting on the papers or showing their support by "voting." The archive is meant to provide permanent site to locate, archive, and share preliminary findings, discussion, and establishing priority. "Nature Precedings is run by Nature Publishing Group (NPG). Nature Publishing Group is a trading name of Macmillan Publishers Ltd, and Nature America Inc., a wholly-owned US subsidiary of Macmillan Publishers Ltd."
4. [PhilPapers](#) is an independent repository and filter site for academic philosophy. Visitors can browse recent journal issues, directly view open access publications, or browse through curated and categorized materials. In addition to locating papers, site members who are "professional authors" (either PhD or already published in an academic journal) can submit papers to the repository, which are made available without peer review. Site members also can comment on individual papers, participate in larger discussions in a categorized forum, or create their own library. All site visitors see suggested related materials and can export citations to reference managers. PhilPapers is based in the UK, and edited by two academics working under an advisory board of six members. The project has in-house developers, and makes its software available as xPapers. Philpapers is sponsored by the Institute of Philosophy School of Advanced Studies, University of London; the UK's Joint Information Systems Committee (Information Environment Programme); and the Centre for Consciousness, Research School of Social Sciences, Australian National University. It is endorsed by The International Association for Computing and Philosophy.
5. [Research Papers in Economics \(RePEc\)](#) is a repository of working papers, journal articles, software, and an index for publications and personnel in the economics field. Institutions and individuals can contribute full-text, abstracts, or links to the database. The database includes an index of published literature, so not all the

material in the database is full-text or available for download in PDF format. Access to the information is free to all visitors. The database is accessible through a number of different websites. All the sites pull from the same database, but not all information is available on all databases. IDEAS, hosted by the Federal Reserve Bank in St. Louis, and EconPapers hosted by the Swedish Business School at Orebro University provide access to the complete database. RePEc is staffed by volunteer “editors” and grew out of the “NetEc group, which received support for its WoPEc project between 1996-1999 by the Joint Information Systems Committee (JISC) of the UK Higher Education Funding Councils, as part of its Electronic Libraries Programme (eLib).” It is unclear how it is financially sustained at present.

6. [science.I/O](#) is a [beta-version] repository for academic work in the computer science and information systems fields. The repository consists of material pulled from arXiv.org, the table of contents from journals and conferences, and submitted by registered users. The content includes conference and journal abstracts, as well as full-text documents. Registered users can submit materials. The materials can be sorted by date or recommendation. There also is an application to build a personalized research feed. The project was founded and developed by a single creator.
7. [The Social Science Research Network \(SSRN\)](#) is a digital repository for abstracts and articles in the social sciences. Visitors can read the latest abstracts provided by journals, publishers, and institutions to the Abstracts Database. Visitors also can read the full text of pre-prints or working papers that individual registered users submitted to the eLibrary. The full text materials are available as PDF downloads. Authors may revise and resubmit later versions of the same paper, and the version history is visible. In addition, the Research Paper Series are topic-based collections by institutes or centers. The Institution Series is specifically for collections provided by institution. SSRN provides article and category-level metrics for downloads, citations, and submissions. SSRN is produced by Social Science Electronic Publishing Inc. (SSEP), an independent, private corporation. They seem to have some revenue from advertising, although it is not clear how they fund the repository.
8. [viXra.org](#) is a repository for e-prints of scientific writing. It allows any viewer to download the materials via PDF. viXra.org currently accepts submissions via email. viXra.org was created to directly compete with arXiv.org because of a disagreement over arXiv.org’s requirement for papers and authors to be endorsed. viXra.org is run by volunteers with donated server space.

Aggregation Projects

1. [Muckrack.com](#) is an aggregation site that collects Twitter posts from journalists. The tweets are presented chronologically, or visitors can browse by news organization, hashtag, topics, or author. The most used and trending hashtags also are visible, as are the latest links included in the tweets. Contextualization, further details about a news event, and informative or administrative content are available on the blog posts. The site is produced by a for-profit company called SawHorse Workshop, which produces a number of websites based on Twitter.
2. [ScienceBlogging Aggregated](#) is a filter blog, or aggregator, of science blogging and writing. Science Blogging Aggregated gathers the content from science publishers around the web through RSS feeds to develop a homepage with the most current information, organized by source. Science Blogging Aggregated also publishes a printed annual anthology of writing from science blogs, called Open Laboratory. Science Blogging Aggregated is a forerunner of ScienceSeeker.
3. [ScienceBlogs](#) is a meta blog of science writing. The site sponsors the sites of more than 80 bloggers to provide content. The featured posts have selected text and a link from the main page. Also provided are direct links to their curated collection of science blogs and RSS feeds. ScienceBlogs is run by a private company and partially funded through advertising.
4. [ScienceSeeker](#) is a filter blog for science blogs that is curated and organized by topic. The titles and links for new posts from member blogs are constantly appearing on ScienceSeeker, categorized by topic, and identified by author. The blogs found on ScienceSeeker either are submitted by the blog authors, other readers, or found by the editors. Human editors categorize the content of each blog. ScienceSeeker is run by the editors without any financial support.

Curation Projects

1. [The Conversation](#) is an Australia-based publisher of original writing that provides academic perspectives on current events, politics, science, and society for a broad audience. Academic authors with detailed credentials provide the content, and trained journalists provide editorial oversight for a consistent writing style across the site. The content is updated twice daily. The academic authors are not paid for their writing, but it is a way for them to prove their public engagement and social impact, which is a required part of their promotion dossier. The Conversation is sponsored through a collaboration of research universities, private companies, and the Australian government.
2. [Journalist's Resource](#) produces and aggregates scholarly resources for professional and student journalists and educators. The website provides links to peer-reviewed policy studies. It also provides reference articles on learning and teaching journalism. All of the materials on the site are curated by the project staff and are attributed to an author. Accompanying links to original studies are an abstract of the research findings and teaching notes. Reference and educational materials are viewable on the site, and these also have an identified author. Material on the site is tagged by topic. Site visitors can comment on the site, but so far it does not seem to be a site for conversation. The project is produced by twelve institutions, and run by the Shorenstein Center on the Press, Politics and Public Policy at Harvard University. It is funded by the Carnegie Corporation and the James L. Knight foundation.
3. [MetaFilter](#) is a community-run filter blog that provides summaries and links to interesting materials around the web. Visitors can read the summaries, follow the links, and read the commentary by community members. Community members can post, link, and comment on site materials. The content is organized by date. The MetaFilter community is restricted to those who purchase an account and meet the expectations of membership. The community has explicit explanations about standards of conduct. There is a \$5 charge to initiate an account. The ability to post is delayed for a few days after registration. Only after providing a few comments is a member allowed to make a post on the main page. MetaFilter is supported by revenue from ads and user registrations.
4. [Phygg.com](#) was an experimental filter website that aggregated and highlighted the papers promoted by visitors to arXiv.org. Viewers read a paper and voted up to move the paper to the main page, or down to move the paper lower on the list. The intended purpose was that visitors to phygg.com could identify the papers considered most valuable to the other readers of arXiv.org. Phygg used a web application named "Pligg" to aggregate the materials. It was produced by a single programmer, which resulted in unreliable site maintenance.
5. [PLoS Biodiversity Hub](#) is a digital project that aggregates and supplements articles from PLoS journals within the theme of biodiversity. Curators select content from PLoS materials, and add value to the articles by providing relevant, supplementary, and publicly available data such as taxonomy, species images, and maps. Site users are able to comment on materials, or share them using social media. PLoS is a non-profit organization supported by philanthropic sources, and revenue from publications fees (APC), advertising, sponsorship, and membership.
6. [Techmeme](#) is a news aggregator focused on the tech industry. Techmeme does not produce original content, but provides links to original content produced by other sources. Techmeme consists of one single page that highlights the most relevant headlines, along with the opening lines, of articles from around the web. The original byline and source appear above the headline, and the viewer follows a link to the original story, or links to other publications covering the same story. The most recent headlines are featured in a sidebar. The site is continuously updated. Software algorithms automatically gather news stories from around the web. Human editors then group and curate the stories according to importance or topic. In July 2010 Techmeme reported 260,000 readers and three million hits per month. The site is run by a private company that employs its editors and is supported by advertising revenue.
7. [Technorati](#) is an aggregator and filter of information from blogs around the web. Technorati functions as a ranker, search engine, and directory for the blogosphere. The site also features original content from member writers. The home page includes summaries, links, and images for a few featured stories written for Technorati. The category tabs at the top link to original content written for Technorati. Below the fold is information pulled from other sources, including a constantly updated list of top blogs, top tags, and the

first few lines of linked blog posts. Technorati provides a ranking of “authority” for blogs through crowdsourcing algorithms: it is determined by the number of links to the site, and the number of times it is referenced by other sites. The authority ranking is in the context of topical categories. The authority, credentials, and sources of its own writers is less clear. Technorati is produced by a private company and employs editors and writers. It is partially supported through ad revenue.

Post-publication Review and Community Discussion

1. [Econ Journal Watch](#) is an online publication that consists of post-publication peer review of articles in economics journals. The reviews are specifically to offer critical reviews rather than highlight important work. Editors referee, and outside reviewers peer review the reviews as well as original essays about the field. The journal is published in editions, available for viewing online and in PDF. Registered users can “talk back” to articles in the journal by writing a comment, and authors of the reviewed materials are given a forum to respond. The journal uses the web application “Journal Talk,” developed specifically for visitor comments. Econ Journal Watch is published by the American Institute for Economic Research and edited by academic economists from George Mason University, the University of Illinois, Chicago, and Montana State University.
2. [Faculty of 1000 \(F1000\)](#) is a digital-only publication, available by subscription, which provides post-publication peer review in the sciences and medicine. Vetted scientists called “Faculty Members” review and draw attention to most interesting articles in their research areas. These reviews are meant to help specialists can stay on top of their field, and help non-specialists quickly identify what is most valuable in related fields. Articles are ranked by the Faculty Member reviewer, and a ranking system for journals is coming soon. F1000 also provides two open access services: a repository for posters, and a journal providing reviews, or “reports” on the trends in research. Faculty Members are selected by nomination, and Junior Faculty Members are sponsored by seniors. F1000 is based in the UK, and is supported through subscription rates tailored to the means of the country in which an institution is based: free for low-income; lower-middle income countries can be sponsored by active Faculty Members. Faculty members do not pay, and authors without an institutional subscription are invited to comment on an evaluation of their article. F1000 is part of the Science Navigation Group, “a group of independent companies that publish and develop information services for the professional biomedical community and the consumer market.”
3. [Jotwell](#) is a website that offers post-publication peer review of legal publishing. The site is divided into topical sections, each with two Section Editors, who oversee ten or more Contributing Editors. Each Contributing Editor provides at least one essay of post-publication peer review that highlights a scholarly work they think deserves attention. Section Editors also approve unsolicited essays for publication. The essays are 500-100 words that summarize and reflect upon the importance of the article under discussion. Site visitors can comment on the essay. All materials are published on the site using WordPress. The project is sponsored and hosted by the University of Miami School of Law.
4. The [polymath blog](#) is a communal project dedicated to solving complicated mathematical problems through collective efforts (called polymath projects). Visitors use commenting tools, a wiki, research threads, and discussion threads to advance the work. Visitors can identify comments that are “useful” or “needs work” in order to solve the problem. Commenting and the Wiki are open to anyone, which makes it a target of spam. The polymath blog uses WordPress and is run by volunteer moderators.
5. [ThirdReviewer](#) (*site currently down*) is a site for post-publication peer review, or “journal club,” for the fields of microbiology, neuroscience, and immunology. ThirdReviewer aggregates abstracts from nearly a dozen major journals, so that comments on the work within the field appear in one venue. ThirdReviewer allows anonymous commenting, and visitors can either “agree” or “disagree” with the comments. ThirdReviewer runs on WordPress and is maintained by one individual who oversees the Neuroscience section, with two others overseeing the Microbiology section. It was active in 2010, but seems to be less active in 2011. They are currently working to incorporate PubMed publications as well.

Options for Traditional Publication on the Open Web

1. [BioMed Central](#) is a digital publisher of peer-reviewed open access science and medicine journals. It began as a free and open publisher, but since its purchase by Springer Science+Business Media, now requires an article processing charge (APC) of \$1685 standard, or \$550 up to \$2570 depending on the journal. BioMed justifies the APC for maintaining and developing open accessibility, tools, and design. A few journals require a subscription access because they “add value to raw data” by commissioning reviews. BioMed journals function like traditional journals in that they have editors, blind peer review, and accept submissions continuously.
2. [OpenEdition](#) is the portal site for the three main programs of The Centre for Open Electronic Publishing (Cléo) that support open source publishing for the humanities and social sciences. Revues.org is the platform is for journals, Hypotheses is the platform for blogs, or “research notebooks,” and Calenda is a centralized calendar for the humanities and social sciences. Journals apply to the Centre for their hosting support. The Centre provides training for the editors of their published journals and blogs. The Centre also provides support to subscribing libraries.
3. [The Public Knowledge Project](#) is an educational center that develops open source software to support the publishing of open access scholarly materials on the internet. These include: Open Journal Systems, Open Conference Systems, Open Harvester Systems, and Open Monograph Press. Open Journal Systems is a locally-installed program that allows editors to manage a journal’s content, submissions, access, indexing, and reading tools. Open Conference System is a locally-installed program that allows editors to create conference websites, collect submissions, allow editing, post conference proceedings, post data sets, register participants, and host online discussions. Open Monograph Press is a locally-installed program to manage the editorial workflow for monographs. The Public Knowledge Project is a partnership of three institutions in Canada and the United States: University of British Columbia, Simon Fraser University, and Stanford University. Their work is funded by Canadian and American government programs, private organizations, and private foundations.
4. [Public Library of Science](#) is a non-profit publisher of open access journals in the sciences. The journals have open submission policies, but the editors oversee the content and peer review process. The articles are available as HTML, PDF, or by reprint when purchased. They are published as final documents, rather than continuously revised documents. PLoS journals provide article metrics, related articles, and commenting and rating functions. PLoS runs on the TOPAZ platform. PLoS relies less on philanthropic support and more on our revenue from publication fees, advertising, sponsorship and membership programs. PLoS charges a publication fee to support the expenses of publication including peer review, journal production, online hosting and archiving. Institutional members are granted discounts on publishing charges, and those from developing countries are not charged. For everyone else, the fees range from \$1350-2900.
5. [Scholarly Exchange](#) is a not-for-profit corporation that provides hosting and support for scholarly journals using the Open Journal Systems platform. Scholarly Exchange provides free hosting, a user-ready installation of Open Journal Systems, and support for the first year of a journal’s publication. Journals are able to choose their domain, language, and customize their editorial process. Each journal can determine its own content, copyright, and revenue policies. The costs of the first year are covered by Scholarly Exchange. After the first year, a charge of \$750 annual fee covers the cost of Scholarly Exchange’s services. Journals also can choose to collect revenue through Google Ads, or pay extra to keep the journal ad-free.