Current status and flow analysis of posted papers in bioRχiv

Hiroyuki Tsunoda[1], Yuan Sun[2], Masaki Nishizawa[3], Xiaomin Liu[4], Kou Amano[5]
[1]Tsurumi University, Japan [2,3]National Institute of Informatics, Japan
[4]National Science Library, Chinese Academy of Sciences, China
[5]National Institute for Materials Science, Japan
[1]tsunoda-h@tsurumi-u.ac.jp [2]yuan@nii.ac.jp [3]nisizawa@nii.ac.jp [4]liuxm@mail.las.ac.cn [5]AMANO.Kou@nims.go.jp

Introduction

The preprint archive is a free online repository and distribution service for unpublished papers. By posting preprints in the archive, authors can make their findings immediately available to the scientific community and receive feedback from readers on draft manuscripts. In recent years, the number of preprint postings has increased significantly. This poster aims to clarify the current status of the bioRχiv archive.

Methods

This study conducted an analysis of papers posted on bioRχiv from November 7, 2013 to February 28, 2019. The web crawler we developed for this purpose visited every page on bioRχiv and downloaded metadata including the title, authors, published digital object identifier (DOI), posted date, and versions of a paper. If the paper was published, it would download the journal title and the published DOI. Next, the metadata were manually verified. Open access journals were defined by adopting an open access publishing model according to Directory of Open Access Journals, Web of Science Selected Open Access by Clarivate Analytics, or the Directory of Open Access Scholarly Resources by International Standard Serial Number International Centre.

<< Top 50 Journals >>

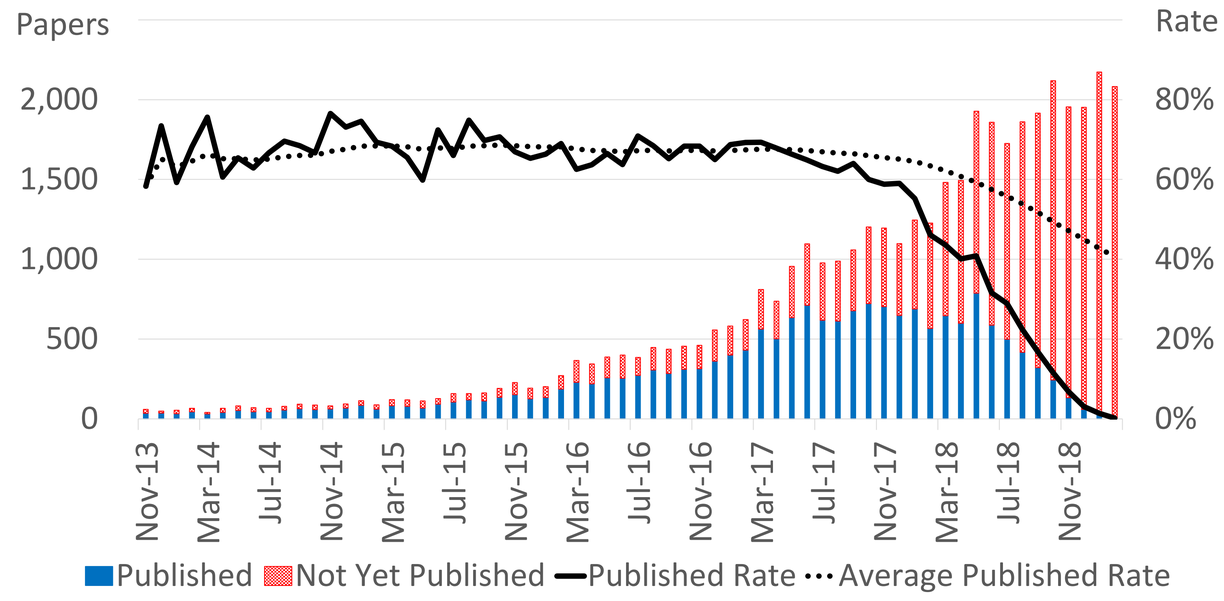
PLoS ONE is an open access journal. As all 902 papers on PLoS ONE are open access, the open paper rate was 100%. On the other hand, Bioinformatics is a hybrid open access journal. Among the 509 papers published in Bioinformatics, 318 are freely available online, thus making the open paper rate 62%.

Journal		Open Paper Rate	
PLoS ONE	OA	902 / 902	100%
Scientific Reports	OA	881 / 881	100%
eLife	OA	866 / 866	100%
Nature Communications	OA	611 / 611	100%
Bioinformatics	HY	318 / 509	62%
PNAS of USA	HY	416 / 487	85%
PLoS Computational Biology	OA	375 / 375	100%
PLoS Genetics	OA	333 / 333	100%
Nucleic Acids Research	OA	285 / 285	100%
Genetics	HY	210 / 283	74%
G3	OA	266 / 266	100%
NeuroImage	HY	36 / 238	15%
BMC Genomics	OA	203 / 203	100%
The Journal of Neuroscience	HY	146 / 197	74%
Genome Biology	OA	191 / 191	100%
Genome Research	HY	184 / 189	97%
Molecular Biology and Evolution	HY	149 / 175	85%
BMC Bioinformatics	OA	157 / 157	100%
Cell Reports	OA	149 / 149	100%
Nature Genetics	HY	5 / 135	4%
PeerJ	OA	132 / 132	100%
mBio	OA	129 / 129	100%
Nature Methods	HY	2 / 126	2%
PLoS Biology	OA	126 / 126	100%
Biophysical Journal	HY	29 / 117	25%
Genome Biology and Evolution	OA	116 / 116	100%
Development	HY	82 / 103	80%
PLoS Pathogens	OA	98 / 98	100%
Gigascience	OA	97 / 97	100%
Molecular Ecology	HY	20 / 94	21%
Molecular Biology of the Cell	HY	52 / 93	56%
Current Biology	HY	28 / 90	31%
Nature	HY	8 / 90	9%
Frontiers in Microbiology	OA	89 / 89	100%
Journal of Virology	HY	51 / 84	61%
Cell	HY	15 / 79	19%
Journal of Theoretical Biology	HY	8 / 79	10%
Journal of Cell Science	HY	56 / 76	74%
Science	HY	21 / 74	28%
American Journal of Human Genetics	HY	31 / 73	42%
Systematic Biology	HY	40 / 73	55%
Cerebral Cortex	HY	29 / 72	40%
Evolution	HY	12 / 70	17%
Journal of Biological chemistry	HY	34 / 70	49%
Journal of Neurophysiology	HY	4 / 70	6%
Neuron	HY	19 / 70	27%
mSphere	OA	69 / 69	100%
The Journal of Cell Biology	HY	65 / 67	97%
Eneuro	NO	0 / 64	0%
F1000Research	NO	0 / 64	0%

Results & Discussion

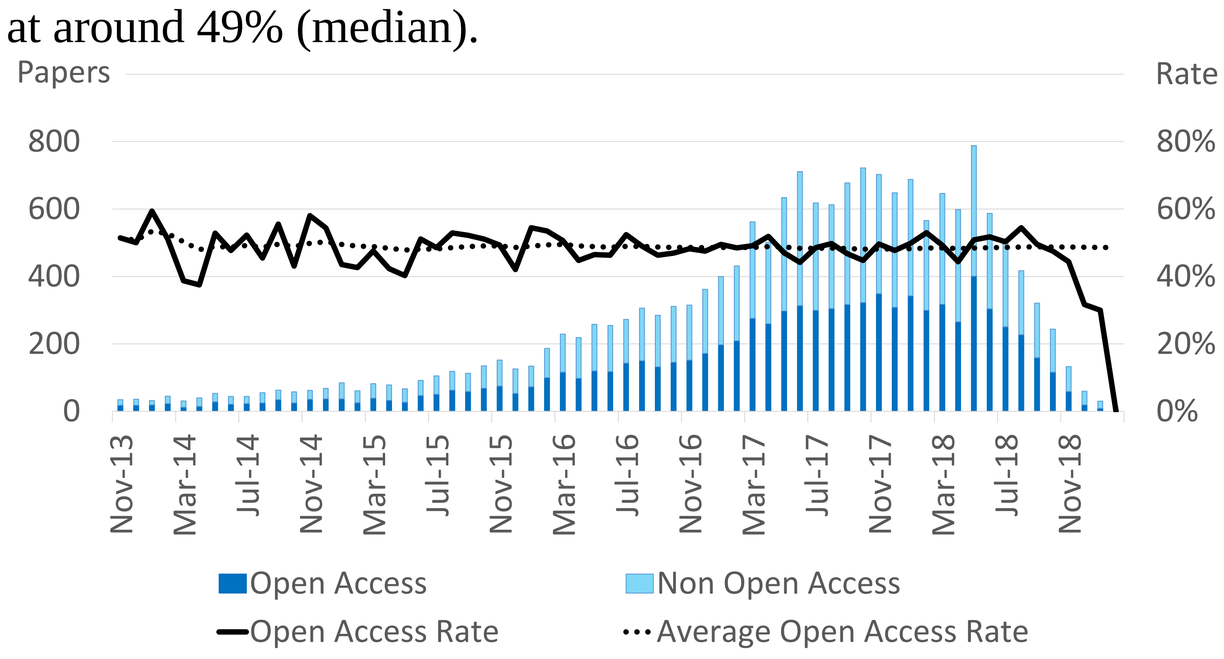
<< Published in Academic Journals >>

43,812 papers had been posted on bioRχiv. 17,818 papers were published in academic journals. The solid line indicates the rate of published papers, and the dotted line indicates the average since its launch. The average publication rate for more than 15 months was uniform at around 67% (median).



<< Open Access Journals>>

Among these, the journals of 8,664 published papers were recorded by the Directory of Open Access Journals, the Web of Science Selected Open Access, or the Directory of Open Access Scholarly Resources. The average open access rate was uniform at around 49% (median).



Hybrid open access journals were defined as those that held papers with a Creative Commons license issued by Impactstory's Unpaywall Database, or the public access papers located on a publisher's site. However, these journals were not listed as Open Access. 12,453 papers were published in open access journals or hybrid open access journals. The average open rate was uniform at around 77% (median).

Conclusion

The conclusions show that the OA model is widely accepted, and has been adopted not only in publishing, but also during the writing process. The concept of "openness" and its practice in academia has inevitably affected the organization of library resources, resource management methods, and service models.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number JP19K12707, JP18K11597, and ROIS NII Open Collaborative Research 2019-(19FS02).