Preprint citation – The Biologist

In 1990, Tim Berners-Lee, an ambitious computer scientist at CERN, built the Web as a way of rapidly sharing information between researchers. Just a few months after, *arXiv* (pronounced 'archive') was developed as a centralised Web-based network for the Maths, Computer Science, and Physics communities. Nearly 30 years on, *arXiv* receives more than 8000 submissions a month of 'pre-prints' – articles that have not yet been formally peer reviewed.

Recently, the Life Sciences seem to have cottoned on to this success, and many researchers are pushing for wider publishing and re-use of pre-prints. Several developments seem to have catalysed this, including *bioRxiv*, an *arXiv* mimic, and the community-led *ASAPbio* initiative that encourages the productive use of pre-prints. Combined with high-level policy changes, such as the NIH now allowing pre-prints in grant submissions, this has resulted in their exceptionally rapid growth in the Life Sciences, now reaching the 1000/month mark.

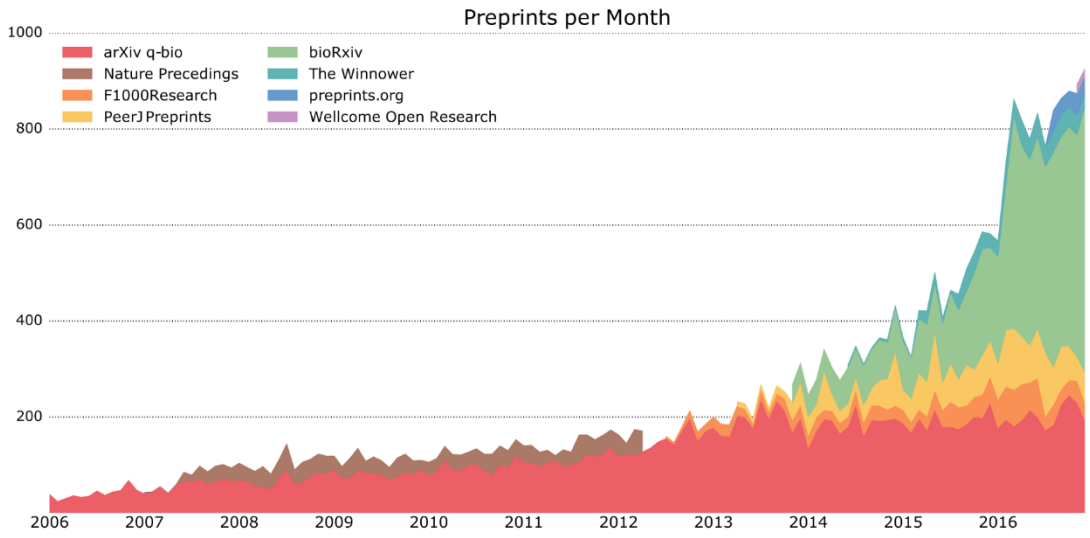

Figure 1: Source data, PrePubMed, courtesy of Jordan Anaya (source)

In spite of this growth, there seems to be a fair amount of resistance to pre-prints. The biosciences are incredibly diverse, including Biochemistry, Zoology, Palaeobiology, and Medicine disciplines, each with their own set of community norms and values. Therefore, it is perfectly understandable that a "one size fits all" solution for pre-prints will never work for the Life Sciences.

One major barrier to pre-print adoption is the question of their citation as scholarly works. Some researchers have claimed that such a practise is inappropriate and constitutes bad scholarship, due to the view that publishing pre-prints is no different to publishing other 'grey literature' due to their 'preliminary' status. This is part of our culture where typically only research that has been explicitly peer reviewed, and therefore has a stamp of certification, is cited.

Such a state is highly different to others, and 4 of the most highly cited 'journals' of all time in Maths and Physics are *arXiv* sub-sections according to Google Scholar. In these communities, a pre-print is considered to be an establishment for priority of that research. However, in Life Sciences, such a status for pre-prints has not been met – they are yet seen to 'count'.

Attempts to close this value gap have largely focussed on making pre-prints more citable from a technical perspective. For example, provision of better metadata, persistent identifiers (DOIs), and even the look and feel more of a traditional journal article. However, this focus seems to have largely missed the point – researchers don't cite pre-prints because it's technically difficult. They don't cite them because they are not deemed worthy of citation. Pre-print resistance has less to do with technology, and more to do with overcoming social barriers.

What researchers rely on are journals (and peer review) to take on the responsibility of telling them what is citable. Pre-prints tell us that the responsibility of the citation lies with the citer, and for some researchers, this is scary. However, what we need to remember is that evaluating the quality and context of research is our job. If we are not doing that then we do not really deserve the title of scholar. There are good and pre-prints, just as there are good and bad papers. As research communities, we should not be using journals as an excuse to absolve ourselves as the ability to think critically.

I recently established *paleorXiv*, a community-led Palaeontology pre-print server. It didn't take long for this to spark a lot of discussion, and I even received an email from a senior researcher emphasising fears that it might be used by creationists to 'get one over' on real science. Yikes! This is a strong sign to those advocating for pre-prints to make sure that we are engaging different communities in their development, and that the social barriers cannot be ignored.

We are still just at the beginning, and there is a long way to go. For *paleorXiv*, we decided to create community-oriented submission guidelines that address many of their concerns, and particularly regarding pre-print citations. To me, the most important is "Please exercise the same care and judgment you would use for any research output when it comes to the citation and re-use of preprints". That's just good scholarly practice.