

Don't be afraid of open data

Björn Brembs 

Published November 16, 2015

Citation

Brembs, B. (2015, November 16). Don't be afraid of open data. *Bjoern.brembs.blog*. <https://doi.org/10.59350/snxxh-06390>

Keywords

Science Politics, Drosophila, Genomics, Open Data

Abstract

This is a response to Dorothy Bishop's post "Who's afraid of open data?". After we had published a paper on how *Drosophila* strains that are referred to by the same name in the literature (Canton S), but came from different [...] [↓ Read the rest of this entry...](#)

Copyright

Copyright © Björn Brembs 2015. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

This is a response to Dorothy Bishop's post "[Who's afraid of open data?](#)".

After we had published a [paper](#) on how *Drosophila* strains that are referred to by the same name in the literature (Canton S), but came from different laboratories behaved completely different in a particular behavioral experiment, [Casey Bergman](#) from Manchester contacted me, asking if we shouldn't sequence the genomes of these five fly strains to find out how they differ. So I went and behaviorally tested each of the strains again, extracted the DNA from the 100 individuals I had just tested and sent the material to him. I also published the behavioral data immediately on our [GitHub project page](#).

Casey then sequenced the strains and made the [sequences](#) available, as well. A few weeks later, both Casey and I were contacted by [Nelson Lau](#) at Brandeis, showing us his bioinformatics analyses of our genome data. Importantly, his analyses wasn't even close to what we had planned. On the contrary, he had looked at something I (not being a bioinformatician) would have considered orthogonal (Casey may disagree). So there we had a large chunk of work we would have never done on the data we hadn't even started analyzing, yet. I was so thrilled! I learned so much from Nelson's work, this was fantastic! Nelson even asked us to be co-authors, to which I quickly protested and suggested, if anything, I might be mentioned in the acknowledgments for "technical assistance" – after all, I had only extracted the DNA.

However, after some back-and-forth, he persuaded me with the argument that he wanted to have us as co-authors to set an example. He wanted to show everyone that sharing data is something that can bring you direct rewards in publications. He wanted us to be co-authors as a reward for posting our data and as incentive for others to let go of their fears and also post their data online. I'm still not quite sure if this fits the [COPE guidelines](#) to the point, but for now I'm willing to take the risk and see what happens.

Nelson is on the tenure clock and so the position of each of his paper's in the journal hierarchy matters. The work is now [online at Nucleic Acids Research](#) and both Casey and I are co-authors. The paper was published before Casey has even gotten around to start his own analyses of our data. This is how science ought to proceed! Now we just need ways to credit such re-use of research data in a currency that's actually worth something and doesn't entail making people 'authors' on publications where they've had little intellectual input. A [modern infrastructure](#) would take care of that...

Until we have such an infrastructure, I hope this story will make others share their data and code as well.