BIOLOGICAL SCIENCES

You Don't Get to 2000 Open Data Sets Without Making a Few Friends – or: How I Got to be Called the Mark Zuckerberg of Open Source Genetics

BASTIAN GRESHAKE

WHAT HAVE I DONE?!

There are many firm believers in the different kinds of openness: open access, open source, open data, open science, open you-name-it. And at least to me, some of the most interesting things happen at the intersection of those different *opens*. Which probably is where *openSNP* – the project I co-founded in 2011 – can be located. It's an open source project which tries to crowdsource collecting open genetic data. This is done by enabling people to donate their personal genetic information into the public domain, alongside phenotypic annotations. And for good measure we also factor in open access, by text mining the *Public Library of Science* and other open databases for primary literature. What started as a somewhat freakish idea in 2011 has by mid–2015 grown to a database of over 2100 genetic data sets and over 300 different phenotypes.

THE IMPACT OF GOING *OPEN*

The existence of *openSNP* is arguably a great success for (citizen) scientists, teachers and everyone else who likes to play around with genetic data. But somewhat ironically, running a project outside academia and designed to overcome the limitations academia has also had lasting impact on my academic career: Besides of learning lots of useful skills due to working on the project, it led to winning awards & grants, finding jobs and collaborators, writing papers, being invited to conferences.

GRANTS & AWARDS

One of the earliest successes was winning the PLOS/Mendeley Binary Battle in 2011, for making the most creative use of their respective open Application Programming Interfaces. This had a twofold effect: First this got us some funding to get the platform up and running and second it also generated wide media coverage, from scientific journals such as Nature to mainstream media like Forbes . Another success grant-wise was winning one of the first *Grants4Apps* by *Bayer HealthCare* in 2013, which again generating some basic funding and media coverage. Speaking of media coverage: it culminated in the probably most ridiculous comparison ever made, the Mark Zuckerberg of open-source genetics.

JOBS & PAPERS

Getting awards and media coverage is nice, but having a job and success in terms of the main academic currency – papers – can be just as important. Being somewhat dragged into the spotlight

additionally helped in those two departments as well. In early 2012 it led to a job as a research assistant at the *Max Planck Institute for Intelligent Systems*, which culminated in the creation of easyGWAS, a web-platform for performing genome-wide association studies in model-organisms and an accompanying manuscript.

Other resulting and still ongoing collaborations are into how Direct-To-Consumer genetic tests can be analyzed using low-budget tools or into how citizen science and participant-led research are challenging the traditional research system. And last but not least, openSNP and the work we've done on it has been published as a open access paper itself with PLOS One.

TRAVEL THE WORLD, MEET INTERESTING PEOPLE

So far, doing open science also gave lots of travel opportunities: Visiting conferences and workshops, in the academic world as well as in industry. Amongst many more, ranging from a panel in San Francisco, over a workshop and Research Hack Days in Zurich to giving talks in Shenzhen and Dublin. While spreading the open science gospel can be very rewarding in itself, meeting others with similar interests can be even more so. Besides finding potential partners in crime (also known as collaborators), it also gives the opportunity to new friendships.

THIS IS THE END

While being an active advocate for openness in academia and science in general is still met with raised eyebrows and shaking heads from time to time, it definitely isn't true any longer that being open necessarily equates career suicide. Publishing open access, doing open source and facilitating open data can also mean a huge career boost, giving networking opportunities, job offers and otherwise unpublished papers. To sum it up: 5/5, would do again.