

What does the open science policy mean for researchers?

Prof. Dr. Jean-Claude Burgelman

Filmed at the Amsterdam Scholarly Summit, 2 July 2019

Why do we do all this? I mean it is because we want better scientists and we want to make, I would say more fun, or interesting for the scientists. What does all this mean for the researcher? I think that we have to train our new generation of researchers with the following new facts of life. So, publications, immediate Open Access. First of all, it will allow you to have your results quicker, and it will be cheaper.

By the way, it has been demonstrated many times that Open Access articles get quoted two to three times more than articles behind paywalls, which means that as a scientist I have two to three times more impact to speak if you follow the logic of the citations.

So, it is an argument we should not forget. The average Open Access book in downloads is a factor up to 10 more downloaded than what you sell, and the price is cheaper to publish an Open Access book. No, things can change of course, but it is undeniable that if you go as a scientist for Open Access, and you leave aside the problems with the impact factor, your science, your results, your impact is bigger. That, I think is a very important argument.

We also need to make sure that our researchers know that data is mandatory, and that they should make data management plans. I'm going to say I find it a little bit strange that sometimes we get criticized on it, but then I say, "Well, okay, whatever research project you do, you produce data because everything is digital, so you must have a plan."

Now, if we can standardize those plans, or you just throw your data in the cloud, no, of course you have a plan. Of course, if you're an astrophysicist or you are researcher in theology, I'm sure that the datasets vary a bit, but that's not the point. The point is that most likely the majority of all science will be extremely data-driven, so you need plans anyway. So, the only thing we try to do is to streamline those plans to make it the standard DMPs required.

I'm absolutely convinced that we will go to a system in the next say 12 to 24 months, that our university associations will come up with a new way to measure all these things, to make it visible, to translate it into indicators, to reward it later on. Most likely, if you go back to the first slide, we see this kind of, now talking let's say five to 10 years ahead, that all research flows in most discipline become called massively online accessible, not free. Open by the way is not free. It depends on the conditions, and then, of course, that all science will be data driven.

So, data science literacy should be, my view, become a mandatory course at least at master's, or at least that at doctoral level. It will also open, by the way, new career paths which we haven't had for the moment. It will also allow open science faster and better publications. We talked about that. Reach of publication opportunities, the whole research value chain is actually to be published, or has already been published and as a result of that you can have a much more broader publication portfolio, which, again, puts a discussion on the impact factor of the single article in a much different perspective.

It falls from all that that we will have a much richer measurement of scientific activity. We might have different career paths. I think that already the National Science Foundation asks for what your five most important products are, not articles, five most important products for the topic of the tenner for which

you are postulating. But that project can be an article, but it can also be an algorithm. It can be a piece of software. It can be a database.

Now, maybe we go to, not maybe, I think we will go to a situation where some of the researchers will not publish articles, but only data sets, or only algorithms, or only visualization software. Why not? If you read the history of statistics, it was also considered as a quite exotic and ridiculous exercise in the '20s, '30s when it started. And nowadays, statistics is some of the most cited articles in the world. So, my guess is based on that historical analogy that we will see the same for data science.

More research potential in the disciplines. It is in particular for social sciences, data science offers an enormous opportunity because it allows you to do things which were never possible before in social sciences, namely having a vast big empirical basis to do your analysis. Most of the social science, you're struck by all the documents, or struck by the fact that you can only do participatory observation for so many people, or via post-doc, or caters. You can try to find out what people thought a week ago.

Now, because of everything becoming a data point, you can actually track social live in real time in a certain sense, right? In Stanford, they have created a Center for Social Physics, and the idea's there is so much data available about social reality that it becomes as empirically solid and robust to study it as physics. A bit of an exacerbation but you see the trend, and we can expand on that later on.

And then, well, more transparency enhances less ethical issues. I would say it is self-evident. If the system is open, well then you simply have less problems. Then I say to research, and you know what? All these tools to do all that, they are freely available. So, this is people in Latin. At the Library of Latin, they have a permanent update of what are all the tools that you can use for publishing, for using licenses, for sharing preprints.

You see all these tools there, and most of it is freely available, so it is not true that this is costly. Most of these are on offer, but there is huge potential, and a huge offer beside the established players if you would want to do so.