

Description

Knowledge should be free to all! Such is the scientific ideal. Free dissemination of information for the benefit of all. The commercial ideal is quite different, of course, more along the line of “knowledge should be secret for our competitive advantage.” Or as it is sometimes phrased, “I’m all in favor of sharing information, so long as it’s them sharing information with us.”

With the vast amounts of data being produced in genetics, meteorology, astronomy, and countless other fields, there is really more than any one researcher can make use of. So why not make the data available to all? More and more scientists are doing just that, and if a few safeguards are taken, I think it’s a good idea.

How to avoid being scooped

Many researchers hesitate [making data available](#) for fear others will use it to [publish papers](#) they were themselves contemplating. This is a valid concern. Unless you have an agreement in place to allow [sharing credit](#) on papers, don’t allow too much information to get out on a work in progress. Some people will take unfair advantage and “scoop” you. This happened to me once and it almost cost my group funding on a project. Fortunately, the story had a happy ending, since the data thieves were unable to follow up with further results on their own and the funding agency turned back to us to finish the project.

When is it time to publish data?

After a mass of data has been mined by yourself for the [lead publications](#), it’s time to make it available to all. Other researchers may be able to make use of the data in ways you never thought of. In my doctoral research, I exploited X-ray crystal data from many papers for purposes I’m sure the authors never considered: it was a small but significant portion of my thesis.

Where to publish and find data?

Some fields of research such as Earth Sciences have data centers for archiving data. Thomas Reuters has a Data Citation Index to facilitate retrieval of data, and funding agencies are increasingly encouraging scientists to share data. If there is no readily available depository, raw data may be included in Supplemental Information sections of a publication.

What’s in it for Me?

Sharing data is not only good for the scientific community; it’s good for a [researcher’s career](#). Studies have shown that papers with data shared in public archives receive more citations than those that hold back such data.

Category

1. Reference Management
2. Reporting Research

Date Created

2014/10/08

Author

daveishan