

| PublisherInfo | | |
|----------------------|---|----------------|
| PublisherName | : | BioMed Central |
| PublisherLocation | : | London |
| PublisherImprintName | : | BioMed Central |

Free access costs money

| ArticleInfo | | |
|-----------------------|---|--|
| ArticleID | : | 4707 |
| ArticleDOI | : | 10.1186/gb-spotlight-20030225-01 |
| ArticleCitationID | : | spotlight-20030225-01 |
| ArticleSequenceNumber | : | 59 |
| ArticleCategory | : | Research news |
| ArticleFirstPage | : | 1 |
| ArticleLastPage | : | 4 |
| ArticleHistory | : | RegistrationDate : 2003-2-25 OnlineDate : 2003-2-25 |
| ArticleCopyright | : | BioMed Central Ltd2003 |
| ArticleGrants | : | |
| ArticleContext | : | 130594411 |

Peg Brickley

Email: pegbrickley@hotmail.com

A recent [National Academies of Science](#)(NAS) report insisting that research data be shared openly was an easy sell to scientists. But convincing funding sources that they should help pay the freight for sharing huge loads of microarray data is not so easy, researchers say.

Released in early February, the NAS study, "[Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Biological Life Sciences](#)," concluded that scientists who publish findings have an ethical duty to allow free and open access to supporting data. Despite a [policy of limiting access](#) to specific "sensitive" data announced last week by a group of journal and author representatives, the principle of open data-sharing remains fundamental, say many researchers for whom the NAS report merely stated the obvious.

"That's a restatement of my own views, and probably the views of a majority of the community," said Gavin Sherlock, director of microarray informatics at [Stanford University](#).

But if the ethics of data sharing are a given, the technology is a work in progress. Unlike genomic data, microarray expression data must be organized and labeled before researchers can work with it and communicate it to other scientists.

"The description of microarray data is quite complicated. It's different from genome sequences which are valuable of themselves. For scientists, it is costly to provide all the necessary meta-information, so depositing it in the database makes sense," said Alvis Brazma, microarray informatics leader at [ArrayExpress](#). One of two public international databases, ArrayExpress is run by the European Bioinformatics Institute (EBI) in the UK.

Brazma is also one of the founding members of the [Microarray Gene Expression Data Society](#) (MGED), which promulgated guidelines for interpreting microarray data called [MIAME](#) (Minimum Information About A Microarray Experiment). Implementing MIAME, however, takes [software and standards](#).

ArrayExpress has developed a Web-based tool that allows scientists to link their data easily to the repository, annotating it in the process. Launched two months ago, the MIAME Express tool has already accepted four complete sets of data and more are on the way.

"This is mostly targeted to smaller laboratories without much programmatic support in-house," Brazma said. The European data bank is putting direct pipelines in place to major centers such as the [Stanford MicroArray Database](#), the [Wellcome Trust Sanger Institute](#) and [The Institute for Genomics Research](#) - connections that capture data automatically while experiments are being done.

Last October, a trio of research journals [adopted the MIAME standard](#) for submitting microarray data for publication, and two of them, *Nature* and *Cell*, went a step further, requiring authors to deposit their data in a public repository as a condition of publication.

But money to support open microarray databases has been scarce on both sides of the Atlantic, and user fees may have to become part of open access to the repositories which both ethics, and now journals, require scientists to use.

The European Union, [European Molecular Biology Laboratory](#) and industry sponsors are paying for a staff of eight people for three years. ArrayExpress has two years of funding left.

Twice, the [National Institutes of Health](#) (NIH) has refused to fund the Stanford Microarray Database, a homegrown effort that may be the largest free and open collection of microarray research data in the country. The National Cancer Institute did provide some early funding, but [Gene Expression Omnibus](#) at the National Center for Biotechnology Information (NCBI) is the NIH-endorsed public database. However, researchers told us that it is slow getting off the ground.

Stanford's self-developed database contains 33,000 micorarrays of information. It took four people working two months to map the attributes of the data in it to the accepted data-exchange module, and it takes eight staffers to maintain the data bank. But the Stanford microarray database serves 85 laboratories on campus, and 400 scientists. The open system has also been adopted freely by other universities, and last year more than a million hits on its Web site were from scientists outside Stanford, Sherlock said.

"We told NIH we think we have an unbelievable service," Sherlock said. "But the review panels said, 'It's not our problem. The PIs doing the research need to provide the database.'"

References

1. National Academies of Science, [<http://www.nas.edu/>]
2. National Research Council Committee on Responsibilities of Authorship in the Biological Life Sciences "Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Biological Life Sciences," February 2003, [http://www.nap.edu/catalog/10613.html?onpi_topnews_020703]
3. P. Park, "New standards for publication of sensitive research," *The Scientist*, February 17, 2003, [<http://www.the-scientist.com/news/20030217/08>]
4. Stanford University, [<http://www.stanford.edu/>]

5. ArrayExpress, European Bioinformatics Institute, [<http://www.ebi.ac.uk/microarray/ArrayExpress/arrayexpress.html>]
6. Microarray Gene Expression Data Society, [<http://www.mged.org/>]
7. MIAME, [<http://www.mged.org/Workgroups/MIAME/miame.html>]
8. L. DeFrancesco, "MIAME begets MAGE," *The Scientist*, September 17, 2002, [<http://www.the-scientist.com/news/20020917/02/>]
9. Stanford MicroArray Database, [<http://www.dnachip.org/>]
10. Wellcome Trust Sanger Institute, [<http://www.sanger.ac.uk/>]
11. The Institute for Genomics Research, [<http://www.tigr.org/>]
12. L. DeFrancesco, "Journal trio embraces MIAME," *The Scientist*, October 10, 2002, [<http://www.the-scientist.com/news/20021010/05/>]
13. European Molecular Biology Laboratory, [<http://www.embl-heidelberg.de/>]
14. National Institutes of Health, [<http://www.nih.gov/>]
15. Gene Expression Omnibus, National Center for Biotechnology Information, [<http://www.ncbi.nlm.nih.gov/geo/>]