



Welcome (Institutional Subscriber : Yale University)

Please login or

TheScientist **monday**

MAGAZINE OF THE LIFE SCIENCES EVERY DAY, ONLINE

- [HOME](#) [CURRENT ISSUE](#) [BROWSE ARCHIVE](#) [MY SCIENTIST](#) [CAREERS](#) [SUBSCRIBE](#) [ADVERTISE](#) [ABOUT THE SCIENTIST](#)

1:51:36 AM

This month:



Volume 20 | Issue 9 | Page 25 | [Reprints](#) | [Issue Contents](#)
[Comment on this article](#)

By Michael Seringhaus and Mark Gerstein

Opinion

The Death of the Scientific Paper

The scientific manuscript as we know it has outlived its usefulness. Here's how to move forward.

- August 2006**
- [Table of Contents](#)
- [Editorial](#)
- [Columns](#)
- [Features](#)
- [Editorial Advisory Board](#)

Today:

The Daily:
[Sign up](#) for The Scientist's daily e-mail.

Blogs:
[Latest Posts](#)

This week:

Podcast:
[TheWeek](#)

Now on iTunes
[subscribe now](#)

For Advertisers:



© RAIMUND KOCH/GETTY



Although the basic currency of science is the research

[Media Kit](#)

[Web Advertising](#)

[Print Advertising](#)

[Contact the Advertising
Department](#)

[Send a digital ad](#)

article, the fruits of modern laboratory research are often incompatible with the aliquot suitable for publication in a scientific manuscript. Genome-scale inquiry and high-throughput experimentation yield enormous data sets, straining the established article framework; meanwhile, isolated findings or negative results are seldom published at all. Further, it has become obvious that preserving data in its native digital format - with search, annotation, and update capabilities - is desirable. Databases are already the primary form of information storage and access for genomics and protein structure research.

The various shortcomings of the article format have been quietly patched with other modes of communication. The typical reader scans general information first - press coverage, textbooks, and high-level descriptions - before exploring in greater detail through PubMed abstracts, conference presentations, and online data sets.

Scientific information is exchanged in a multi-tiered manner, and these myriad other channels render the scientific manuscript optional, if not obsolete. For instance, those seeking authoritative high-level scientific knowledge can visit the NCBI Bookshelf, an indexed and fully searchable digital archive of textbooks with citations linking directly to PubMed abstracts; a scientist in search of genomic data or bioinformatics software need look no further than online databases or laboratory Web sites. Often the journal article, the bedrock of peer-reviewed scientific knowledge, is the last information source consulted.

While this highlights the importance of nontraditional communication in science, it is also regrettable: After all, journal articles are the main output for which scientists earn recognition, and producing them commands a huge share of our efforts. Meanwhile, virtually no credit is afforded to producing quality high-level summaries or to online data deposition.

Journals must produce more than just papers. Editors should demand online deposit of data as a requirement for publication, and enforce a unified nomenclature for biology. In addition to the traditional manuscript,

authors should deliver structured methods and results sections suitable for computer parsing, a lay-friendly news blurb (like those *PLoS Medicine* includes), and a single PowerPoint slide summarizing the work. This entire body of information should be peer-reviewed, published en masse, and kept in sync, thereby avoiding the current problem of disconjugate articles and data sets.

Broadly, such publishing reform would expand the purview of journals to other tiers of scientific content. The next step is to consolidate all tiers into a single searchable resource. We envision a centralized digital index acting on all information in the biomedical sciences. Just as PubMed indexes journal abstracts in a structured fashion, we propose cataloging a broad range of material, which would enable users to run PubMed-like queries over abstracts, full text, data sets, lay summaries, and presentations, all through a single portal.

Of course, to some degree this goal mimics that of existing entities such as the NCBI's Entrez. The major difference is that the NCBI approach is monolithic: an attempt to amass and house all scientific communication in one place. This is neither realistic nor desirable. We must recognize the plurality of voices contributing to science worldwide. The driving force behind data integration should not be a single American entity; instead, it should be a collaborative effort driven by journals: decentralized information, central access.

This central index would add value by cataloging and interrelating disparate data sources. For instance, a data set might link not only to its companion article, but also to earlier versions of the data, news coverage, reviews, and related talks given by the authors. Community annotation and discussion would add another dimension to peer review, and interested parties of all pedigrees could access information at a level suitable to their needs.

The future of scientific data lies in digital storage and access. It makes sense to revamp academic publishing now to ensure efficient database deposit. Today, considerable resources are poured into extracting data

from journal articles; indeed, many databases are still hand-curated by dedicated staff. There will be some up-front costs to implementing this system, but a transition to include machine-readable output will soon pay for itself. Forget "publish or perish." Academic publishing must diversify or die.

Michael Seringhaus is a graduate student in Mark Gerstein's group at Yale University, where Gerstein is A. L. Williams Professor of Biomedical Informatics.

mseringhaus@the-scientist.com

[Comment on this article](#)

© 1986-2006 *The Scientist*



sponsored links